



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 212039

TO: Manjunath N Rao
Location: REM/2A01/2C70
Art Unit: 1652
Monday, January 08, 2007
Case Serial Number: 10/800684

From: Kristine Hensle
Location: Biotech-Chem Library
REM-1A70
Phone: (571)272-4161

Kristine.Hensle@uspto.gov

Search Notes

Examiner Rao,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Kristine Hensle
Reference Librarian
STIC Biotech/Chem Library
(571)272-4161

This Page Blank (uspto)

STIC-Biotech/ChemLib

From: Chan, Christina
Sent: Friday, January 05, 2007 4:43 PM
To: Rao, Manjunath N.; STIC-Biotech/ChemLib
Subject: RE: Request for RUSH sequence search for 10/800684

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Rao, Manjunath N.
Sent: Friday, January 05, 2007 2:48 PM
To: Chan, Christina
Subject: Request for RUSH sequence search for 10/800684

Hello Christina,

Please authorize the request below as RUSH. The reason being this is an amended case that is due in the next bi-week.

Thanks
-Manjunath

Manjunath N. Rao, Ph.D.
Primary Patent Examiner
Art Unit 1652, Room 2A01
Remsen Building
United States Patent and Trademark Office
400, Dulany St, VA

Phone: 571-272-0939
Fax: 571-273-0939

From: Manjunath N. Rao
Art Unit 1652, Room 2A01
Mail Box in Room 2C70
Phone: 272-0939

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

This Page Blank (uspto)

Date: 1-5-07

Please search the following as soon as possible for application with serial number

10/800684

1. SEQ ID NO: 1, and 5 against all commercial amino acid sequence databases, issued patents/published applications amino acid sequence database and pending application amino acid sequence database. Please provide a print of results.

If you have any questions please call me at the above phone number.

Thanks

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

This Page Blank (uspto)

STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor
571-272-2507 Remsen 1 A51

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library Remsen Bldg.

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2007 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2007, 22:19:02 ; Search time 142 Seconds
(without alignments)
2121.871 Million cell updates/sec

Title: US-10-800-684-1
Perfect score: 3437
Sequence: 1 MKRLCAVLTALVGLAGT.....YAYSTYGMADYQLKAVVYTG 659

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq.8:*
1: geneeqp1980s:*
2: geneeqp1990s:*
3: geneeqp2000s:*
4: geneeqp2001s:*
5: geneeqp2002s:*
6: geneeqp2003as:*
7: geneeqp2003bs:*
8: geneeqp2004s:*
9: geneeqp2005s:*
10: geneeqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------|--------------------|
| 1 | 3437 | 100.0 | 659 | AAW24121 | Thermococ |
| 2 | 3437 | 100.0 | 659 | AAW94840 | AAW94840 W09856926 |
| 3 | 2914 | 84.8 | 659 | AAW24123 | Protease. |
| 4 | 2664 | 77.5 | 663 | ADN46373 | Thermococ |
| 5 | 2483.5 | 72.3 | 654 | AAW24129 | Pyrococcu |
| 6 | 2483.5 | 72.3 | 654 | AAW94841 | Hyperther |
| 7 | 2138.5 | 62.2 | 522 | AAW24122 | Pyrococcu |
| 8 | 2138.5 | 62.2 | 522 | AAW94838 | Hyperther |
| 9 | 1707 | 49.7 | 412 | AAW94836 | Hyperther |
| 10 | 1154 | 33.6 | 237 | AAW87009 | Peptide s |
| 11 | 935 | 27.2 | 188 | AAW87011 | Peptide s |
| 12 | 935 | 27.2 | 188 | AAW24126 | Thermococ |
| 13 | 918.5 | 26.7 | 545 | ABW09483 | T. yonsei |
| 14 | 665 | 19.3 | 734 | AAW09483 | Streptomy |
| 15 | 665 | 19.3 | 823 | AAW13667 | AAW13667 |
| 16 | 647.5 | 18.8 | 520 | AAW13668 | Dhpa-mel |
| 17 | 636 | 18.5 | 1079 | AAW13666 | Fragment |
| 18 | 636 | 18.5 | 1079 | ABW81180 | Transglut |
| 19 | 632.5 | 18.4 | 443 | ABW07391 | Foreign p |
| 20 | 632.5 | 18.4 | 443 | ADM99125 | Bacterial |
| 21 | 590.5 | 17.9 | 1237 | ABW11343 | Protein e |
| 22 | 590.5 | 17.2 | 182 | ADM99123 | Bacterial |
| 23 | 577.5 | 16.8 | 168 | ABW09484 | T. yonsei |
| | | | | AAW24127 | Pyrococcu |

| | | | | | | |
|----|-------|------|-----|----|----------|-----------|
| 24 | 576.5 | 16.8 | 552 | 8 | ADM99241 | Environme |
| 25 | 565.5 | 16.5 | 383 | 5 | ADW54158 | L-25 Kera |
| 26 | 564.5 | 16.4 | 379 | 6 | AAE29944 | Bacillus |
| 27 | 563.5 | 16.4 | 379 | 10 | AEF11989 | Bacillus |
| 28 | 561.5 | 16.3 | 379 | 2 | AAW04137 | Sequence |
| 29 | 560.5 | 16.3 | 379 | 2 | AAW88521 | Keratinas |
| 30 | 559.5 | 16.3 | 379 | 5 | AAW26762 | Bacillus |
| 31 | 559.5 | 16.3 | 379 | 2 | AAE19062 | Bacillus |
| 32 | 557.5 | 16.2 | 350 | 2 | AAW46598 | Amno aci |
| 33 | 557.5 | 16.2 | 379 | 2 | AAW31600 | Subtilist |
| 34 | 557.5 | 16.2 | 379 | 6 | AAE29943 | Bacillus |
| 35 | 557.5 | 16.2 | 379 | 7 | ADF70896 | Subtilist |
| 36 | 557.5 | 16.2 | 379 | 8 | ADH39841 | B. amylo |
| 37 | 557.5 | 16.2 | 379 | 8 | ADQ58783 | B. lichen |
| 38 | 557.5 | 16.2 | 379 | 8 | ADU64368 | Bacillus |
| 39 | 557.5 | 16.2 | 380 | 8 | ADH39843 | Subtilist |
| 40 | 556.5 | 16.2 | 379 | 6 | AAE29945 | Bacillus |
| 41 | 555.5 | 16.2 | 379 | 8 | ADQ58791 | B. lichen |
| 42 | 554.5 | 16.1 | 372 | 6 | AAE29942 | Bacillus |
| 43 | 554.5 | 16.1 | 379 | 8 | ADQ58786 | B. lichen |
| 44 | 551.5 | 16.0 | 378 | 6 | AAE29947 | Bacillus |
| 45 | 550.5 | 16.0 | 379 | 6 | AAE29946 | Bacillus |

ALIGNMENTS

| | | |
|-------------|---|----------------------------|
| RESULT 1 | AAW24121 | standard; protein; 659 AA. |
| ID | AAW24121 | |
| XX | AAW24121; | |
| AC | 17-OCT-2003 (revised) | |
| DT | 20-APR-1998 (first entry) | |
| DT | 20-APR-1998 | |
| XX | Thermococcus protease. | |
| DE | Thermococcus protease. | |
| XX | Protease; research reagent; thermal stability; thermococcus celer. | |
| XX | Protease; research reagent; thermal stability; thermococcus celer. | |
| OS | Thermococcus celer; DSM-2476. | |
| XX | Thermococcus celer; DSM-2476. | |
| XX | WO9721823-A1. | |
| PN | WO9721823-A1. | |
| XX | 19-JUN-1997. | |
| PD | 19-JUN-1997. | |
| XX | 07-NOV-1996; 96WO-JP003253. | |
| PF | 07-NOV-1996; 96WO-JP003253. | |
| XX | 12-DEC-1995; 95JP-00323285. | |
| PR | 12-DEC-1995; 95JP-00323285. | |
| XX | (TAKI) TAKARA SHUZO CO LTD. | |
| PA | (TAKI) TAKARA SHUZO CO LTD. | |
| XX | Takakura H, Morishita M, Yamamoto K, Mitra M, Asada K; | |
| XX | Tsunasawa S, Kato I; | |
| PI | WPI: 1997-332794/30. | |
| DR | N-PSDB; AAT85667. | |
| XX | Protease(s) and genes encoding them obtained from Thermococcus and | |
| PT | Pyrococcus strains - have extremely high thermal stability and are useful | |
| PT | industrially and as research reagents. | |
| XX | Claim 1; Page 82-85; 159pp; Japanese. | |
| PS | This sequence represents the protease from Thermococcus celer DSM-2476. | |
| XX | This sequence is a protease of the invention. The proteases of the | |
| CC | invention have extremely high thermal stability. The proteases can be | |
| CC | used as research reagents, and industrially in the food, drug and | |
| CC | chemical industries. (Updated on 17-OCT-2003 to standardise OS field) | |
| XX | Sequence 659 AA; | |
| SQ | Sequence 659 AA; | |
| Query Match | 100.0%; Score 3437; DB 2; Length 659; | |

Best Local Similarity 100.0%; Pred. No. 5.7e-203;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MKRLGAVVVALVVLVGLAGTALAAPVKPYVRNNAVOOKKYGILLTGLEFKVQRMMNNEV 60
DB 1 MKRLGAVVVALVVLVGLAGTALAAPVKPYVRNNAVOOKKYGILLTGLEFKVQRMMNNEV 60
QY 61 DVIYMGSGYGDNDRAVKVRLMGAQVKSYSKIIIPAAVKIKARDLLIAGMIDTGFQNT 120
DB 61 DVIYMGSGYGDNDRAVKVRLMGAQVKSYSKIIIPAAVKIKARDLLIAGMIDTGFQNT 120
QY 121 RVSGIKFIEDYKVOVDATSVSQIGADTVNMSLGSDSGVVAIVDTGIDANHPDLKXK 180
DB 121 RVSGIKFIEDYKVOVDATSVSQIGADTVNMSLGSDSGVVAIVDTGIDANHPDLKXK 180
QY 181 VIGWYDAVNGRSTPYDDQGHGTHVAGIYAGTGSVNSQYIGVAFPAKLVGVVVLGADSGS 240
DB 181 VIGWYDAVNGRSTPYDDQGHGTHVAGIYAGTGSVNSQYIGVAFPAKLVGVVVLGADSGS 240
QY 241 VSTIIAGVDMVQNDKXGIRVINISLSSSSQSDGTDLSQAANNAMDAIIVCVAAGNS 300
DB 241 VSTIIAGVDMVQNDKXGIRVINISLSSSSQSDGTDLSQAANNAMDAIIVCVAAGNS 300
QY 301 GPNITVGSPPAAASKVITVGAVDSDNDIASFSSRGPTADGRLKPEVAPGVDTIAPRAGS 360
DB 301 GPNITVGSPPAAASKVITVGAVDSDNDIASFSSRGPTADGRLKPEVAPGVDTIAPRAGS 360
QY 361 TSMGTPINDYTTKASGTSMATPHVSGVALILQAHPSMTDPKVTALILETADIAPRKEIA 420
DB 361 TSMGTPINDYTTKASGTSMATPHVSGVALILQAHPSMTDPKVTALILETADIAPRKEIA 420
QY 421 DIAYGAGRVNYYKAIKDDVAKLFTFGSVADKSGATFTFYSGATFTYATILYMDTSSDI 480
DB 421 DIAYGAGRVNYYKAIKDDVAKLFTFGSVADKSGATFTFYSGATFTYATILYMDTSSDI 480
QY 481 DLYLYDPNGNEVDYSYTAAYGFEKVGYYNPAGTWTKVVSYKGAANYQVDVSDGSLQ 540
DB 481 DLYLYDPNGNEVDYSYTAAYGFEKVGYYNPAGTWTKVVSYKGAANYQVDVSDGSLQ 540
QY 541 SGGGNPNPNPNPNPTPTTDQTFPGSVNDYWDTSDFPTMNVNSGATKITGDLTFDTSYND 600
DB 541 SGGGNPNPNPNPNPTPTTDQTFPGSVNDYWDTSDFPTMNVNSGATKITGDLTFDTSYND 600
QY 601 LDLYLYDPNGNLVDNSTSSNSYEHVEYANPAPGWTFTFLVYASTYGADVQLAAVYYG 659
DB 601 LDLYLYDPNGNLVDNSTSSNSYEHVEYANPAPGWTFTFLVYASTYGADVQLAAVYYG 659

```

RESULT 2

AAM94840 standard; protein; 659 AA.

```

XX AAM94840;
AC AAM94840;
DT 06-MAY-1999 (first entry)
DE WO9856926 Seq ID 12.
XX
XX Hypothermostable; protease; thermophilic; bacterium; subtilisin;
XX additive; drug; washing agent; foodstuff; chemical synthesis.
OS Thermococcus celer.
XX
XX WO9856926-A1.
XX
XX 17-DEC-1998.
XX
XX 04-JUN-1998; 98WO-JP002465.
XX
XX 10-JUN-1997; 97JP-00151969.
XX
XX (TAKI ) TAKARA SHUZO CO LTD.
XX

```

PI Takakura H, Morishita M, Shimoto T, Asada K, Kato I;
DR WPI; 1999-080907/07.
XX N-PSDB; AAX05926.
XX
XX Recombinant hyperthermostable protease from *Pyrococcus furiosus* - and
PT gene encoding it, for large scale production of the protease for
XX industrial use.
XX
XX Disclosure; Page 55-58; 82pp; Japanese.

Sequence 659 AA:

Query Match 100.0%; Score 3437; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 5.7e-203;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MKRLGAVVVALVVLVGLAGTALAAPVKPYVRNNAVOOKKYGILLTGLEFKVQRMMNNEV 60
DB 1 MKRLGAVVVALVVLVGLAGTALAAPVKPYVRNNAVOOKKYGILLTGLEFKVQRMMNNEV 60
QY 61 DVIYMGSGYGDNDRAVKVRLMGAQVKSYSKIIIPAAVKIKARDLLIAGMIDTGFQNT 120
DB 61 DVIYMGSGYGDNDRAVKVRLMGAQVKSYSKIIIPAAVKIKARDLLIAGMIDTGFQNT 120
QY 121 RVSGIKFIEDYKVOVDATSVSQIGADTVNMSLGSDSGVVAIVDTGIDANHPDLKXK 180
DB 121 RVSGIKFIEDYKVOVDATSVSQIGADTVNMSLGSDSGVVAIVDTGIDANHPDLKXK 180
QY 181 VIGWYDAVNGRSTPYDDQGHGTHVAGIYAGTGSVNSQYIGVAFPAKLVGVVVLGADSGS 240
DB 181 VIGWYDAVNGRSTPYDDQGHGTHVAGIYAGTGSVNSQYIGVAFPAKLVGVVVLGADSGS 240
QY 241 VSTIIAGVDMVQNDKXGIRVINISLSSSSQSDGTDLSQAANNAMDAIIVCVAAGNS 300
DB 241 VSTIIAGVDMVQNDKXGIRVINISLSSSSQSDGTDLSQAANNAMDAIIVCVAAGNS 300
QY 301 GPNITVGSPPAAASKVITVGAVDSDNDIASFSSRGPTADGRLKPEVAPGVDTIAPRAGS 360
DB 301 GPNITVGSPPAAASKVITVGAVDSDNDIASFSSRGPTADGRLKPEVAPGVDTIAPRAGS 360
QY 361 TSMGTPINDYTTKASGTSMATPHVSGVALILQAHPSMTDPKVTALILETADIAPRKEIA 420
DB 361 TSMGTPINDYTTKASGTSMATPHVSGVALILQAHPSMTDPKVTALILETADIAPRKEIA 420
QY 421 DIAYGAGRVNYYKAIKDDVAKLFTFGSVADKSGATFTFYSGATFTYATILYMDTSSDI 480
DB 421 DIAYGAGRVNYYKAIKDDVAKLFTFGSVADKSGATFTFYSGATFTYATILYMDTSSDI 480
QY 481 DLYLYDPNGNEVDYSYTAAYGFEKVGYYNPAGTWTKVVSYKGAANYQVDVSDGSLQ 540
DB 481 DLYLYDPNGNEVDYSYTAAYGFEKVGYYNPAGTWTKVVSYKGAANYQVDVSDGSLQ 540
QY 541 SGGGNPNPNPNPNPTPTTDQTFPGSVNDYWDTSDFPTMNVNSGATKITGDLTFDTSYND 600
DB 541 SGGGNPNPNPNPNPTPTTDQTFPGSVNDYWDTSDFPTMNVNSGATKITGDLTFDTSYND 600
QY 601 LDLYLYDPNGNLVDNSTSSNSYEHVEYANPAPGWTFTFLVYASTYGADVQLAAVYYG 659
DB 601 LDLYLYDPNGNLVDNSTSSNSYEHVEYANPAPGWTFTFLVYASTYGADVQLAAVYYG 659

```


CC inspection, molecular biology and immunology. With this method, the
CC disruption of a gene at an arbitrary position in a genome can be achieved
CC efficiently and reliably. The present sequence is that of a protein
CC encoded by the genome of *Thermococcus kodakarensis* which was derived
CC using the method of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_sequences

XX Sequence 663 AA;

Query Match 77.5%; Score 2664; DB 8; Length 663;

Best Local Similarity 75.9%; Pred. No. 2.1e-155;
Matches 505; Conservative 64; Mismatches 88; Indels 8; Gaps 4;

```
QY 1 MKRLGAVVLAIVLVGLAGTALAAPVKEPVVRNNAVQKRGILLTPGLFKKVGQNMNQEV 60
DB 1 MKKFGAVVLAFLVGLMAGSVLAAPQKPAVR-NVSOQKNYGLLTPGLFKKVGQNMNQEV 59
QY 61 DTIVIMFGSYGDRDAVAVKVLMLGQVYKSKITPAVAVKIKARDLLIAGMIDTGYRGNT 120
DB 60 STIIMFNQADKEKAVBILDFLGAKIKYNIHIIIPALAVKIKVLDLIIAGLMDTGIFGNA 119
QY 121 RVSGIKFIQEDYKYQVDDAT-----SVSQIADTVNMSLGVDGSGVVAIVDTGIDANHP 175
DB 120 QLSGVQFIQEDYKYKAVETBGLDESAQVWATMMN-LGYDSGIGITIGIIDGIDASHP 178
QY 176 DLKRGVIGWYDAVNGRSTPYDDQGHGTHVAGIVAGTGV-NSQYIGVAPGAKLVGVVVG 234
DB 179 DLQKGVIGWYDFVNGKTTPTDYDNGHGVASIAAGTAAANGKKGAPAKLVGIVLVN 238
QY 235 ADGSGYSTIIAGVWVYVONKDKGIRVYNISLGSSQSSDGTDSLQAVNNAMDAIGVVC 294
DB 239 GQSGSISDIIINGVDMVONKDKGIRVYNISLGSSQSSDGTDSLQAVNNAMDAIGLVV 298
QY 295 VVAAAGNSGPNNTYVGSPPAAASKVITVGAVDSNDNIASFSSRGPADGRLEKEVVA PGVDI 354
DB 299 VVAAAGNSGPNKTYVGSPPAAASKVITVGAVDKVDVITDSSRGPADNRLEKEVVA PGWMI 358
QY 355 APRAAGTSMGTPINDYTTKASGTSMAITPHVSGVGLLIQAHPSMTPDKVKTALILETDIV 414
DB 359 AARASGTSMGQPIINDYTTAAPTSMATPHVAGIAALLQAHPSMTPDKVKTALILETDIV 418
QY 415 APKEIADIAYAGRVNVYKAIKYDDYAKLFTGSAADKGSATHTFEDVSGATFVATLYMD 474
DB 419 KPDEIADIAYAGRVNAKAAIYDNYAKLFTGVSNGKSGSHQPTTISGAGFVATILYMD 478
QY 475 TGSSDIDLXYLDPNGNEVDYSYTAAYGFEKGYNNPTAGTWTVKVSYKGAANYQVDVVS 534
DB 479 NSGSDLDLXYLDPNGNQVDYGTAYGFEKGYNNPTAGTWTIKVSYSGSANYQVDVVS 538
QY 535 DGSLSQSGGQGNPNPNPNPTPTDTQTFTTGSVNDYMDTSPTFMNNSGATKLTGDLTF 594
DB 539 DGSLSQSGGQSGSPSPSPPEPTDEKFTGTGVHDYDKSDTFTMTVNSGATKLTGDLTF 598
QY 595 DTSVNDLXYLYDPNGNIVDRSTSSNSYEHVEYANPAPGWTFLYYASTGMADYOLKA 654
DB 599 DTSVNDLXYLYDPNQNLVDRSSNSYEHVEYNNPAPGWTFLYYAVDTGTGYADYOLDA 658
QY 655 VVYYG 659
DB 659 KVVYV 663
```

RESULT 5
AAW24129
AAW24129 standard; protein; 654 AA.XX AAW24129;
AC
DT 17-OCT-2003 (revised)
DT 20-APR-1998 (first entry)
XXDE *Pyrococcus furiosus* protease PFUS.XX Protease; research reagent; thermal stability; *pyrococcus furiosus*.XX *Pyrococcus furiosus*; DSM-3638.

XX WO9721823-A1.

XX 19-JUN-1997.

XX 07-NOV-1996; 96WO-UP003253.

XX 12-DEC-1995; 95UP-00323285.

XX (TAKI) TAKARA SHUZO CO LTD.

XX Takakura H, Moriehita M, Yamamoto K, Mita M, Asada K;

XX Tsunashima S, Kato I;

XX MPI; 1997-332794/30.

XX N-PSDB; AAT85695.

XX Protease(s) and genes encoding them obtained from *Thermococcus* and
XX *Pyrococcus* strains - have extremely high thermal stability and are useful
XX industrially and as research reagents.

XX Disclosure; Page 125-128; 159pp; Japanese.

XX This sequence represents the protease from *Pyrococcus furiosus* DSM-3638.XX This sequence is a protease of the invention. The proteases of the
XX invention have extremely high thermal stability. The proteases can be
XX used as research reagents, and industrially in the food, drug and
XX chemical industries. (Updated on 17-OCT-2003 to standardise OS field)

SQ Sequence 654 AA;

Query Match 72.3%; Score 2483.5; DB 2; Length 654;

Best Local Similarity 72.1%; Pred. No. 2.6e-144;
Matches 481; Conservative 69; Mismatches 96; Indels 21; Gaps 8;

```
QY 1 MKRLGAVVLAIVLVGLAGTALAAPVKEPV--VRNNAVQKRGILLTPGLFKKVGQNMNQ 58
DB 1 MKRLKALIVLVGLAGVGSVVAAPKKEVQVRN--VEKNIGLLTPGLFKKIQKLNPNH 57
QY 59 EVDTVIMFGSYGDRDAVAVKVLMLGAQVYKSKITPAVAVKIKARDLLIAGMIDTGYRG 118
DB 58 EISTVIVFENHREKELAVRYLEMGAKRVYVHIIPALADLKVRLDIVSGD--TG--G 113
QY 119 NTVSGIKFIQEDYKYQVDDA-----TSVQIADTVNMSLGVDGSGVVAIVDTGIDAN 173
DB 114 KAKISGVRFIQEDYKYVSALEBGLDESAQVWATVMN-LGYDSGIGITIGIIDGIDAS 172
QY 174 HPDLKGVIGWYDAVNGRSTPYDDQGHGTHVAGIVAGTGV-NSQYIGVAPGAKLVGVV 232
DB 173 HPDLQGVIGWYDFVNGRSTPYDDHGHGVASIAAGTGAASNGKKGAPAKLVGIV 232
QY 233 LGADGSGYSTIIAGVWVYVONKDKGIRVYNISLGSSQSSDGTDSLQAVNNAMDAIGV 292
DB 233 LGADGSSISITIIKGYEMAVDNDKDKGIRVYNISLGSSQSSDGTDSLQAVNNAMDAIGV 292
QY 293 VCVAAAGNSGPNNTYVGSPPAAASKVITVGAVDSNDNIASFSSRGPADGRLEKEVVA PGVD 352
DB 293 VVVAAGNSGPNKTYIGSPAAASKVITVGAVDKVDVITDSSRGPADNRLEKEVVA PGW 352
QY 353 IIAPRASGTSMGTPINDYTTKASGTSMAITPHVSGVGLLIQAHPSMTPDKVKTALILETD 412
DB 353 IIAPRASGTSMGQPIINDYTTAAPTSMATPHVAGIAALLQAHPSMTPDKVKTALILETD 412
QY 413 IVAPEIADIAYAGRVNVYKAIKYDDYAKLFTGSAADKGSATHTFEDVSGATFVATLY 472
DB 413 IVKPEIADIAYAGRVNAKAAIYDNYAKLFTGVYANKSGQTHGVISGASFTVATLY 472
QY 473 WDTGSSDIDLXYLDPNGNEVDYSYTAAYGFEKGYNNPTAGTWTVKVSYKGAANYQVDV 532
```


| | | | |
|----------|---|---|-----|
| Db | 473 | MDNANSDDDLVLYDPNGNGVQVDSTAYAYGFEAKYGVNPNFDGKMTIKVYSYSGSANQVDV | 532 |
| Qy | 533 | VSDGSLGSGGGGANGPNPNPNFTPTTDIOTFFGSVNDYWDTSDFETNANSGATKITGD | 592 |
| Db | 533 | VSDSLSLG-----PGSSPSPOPEPTVDAKTFQSGSHYYVDRDPTFTMTNNGATKITGDL | 587 |
| Qy | 593 | TEPDSYNDLDDLYLYXPNNGNLVDRSTSSNSYEHVEYANPARGMTPLVVAYSTYGMADYOL | 652 |
| Db | 588 | VFDTSYHDLDDLYLYDPNQKLVDRESPPNSYEHVEYILTPAGIMYFLVVAIYTYGMAYTEL | 647 |
| Qy | 653 | KAVVYYG 659 | |
| Db | 648 | TAKVYYG 654 | |
| RESULT 6 | | | |
| ID | AAM94841 | standard; protein; 654 AA. | |
| XX | AAM94841; | | |
| XX | 06-MAY-1999 | (first entry) | |
| XX | DE | Hyperthermostable protease. | |
| XX | KM | Hyperthermostable; protease; thermophilic; bacterium; subtilisin; | |
| XX | KM | additive; drug; washing agent; foodstuff; chemical synthesis. | |
| OS | Pyrococcus furiosus. | | |
| XX | MO9856926-A1. | | |
| PN | 17-DEC-1998. | | |
| PD | 04-JUN-1998; | 98WO-JP002465. | |
| PF | 10-JUN-1997; | 97JP-00151969. | |
| PR | (TAKI) TAKARA SHUZO CO LTD. | | |
| XX | Takakura H, Morishita M, Shimojo T, Asada K, Kato I; | | |
| PI | WPI; 1999-080907/07. | | |
| DR | N-PSDB; AAK05929. | | |
| DR | Recombinant hyperthermostable protease from <i>Pyrococcus furiosus</i> - and | | |
| PT | gene encoding it, for large scale production of the protease for | | |
| PT | industrial use. | | |
| PS | Disclosure; Page 60-63; 82pp; Japanese. | | |
| XX | The invention relates to a hyperthermostable protease derived from a | | |
| CC | thermophilic bacterium (especially <i>Pyrococcus furiosus</i>). The protease has | | |
| CC | working temperature 40-110 deg.C (optimum 80-95 deg.C), working pH 5-10 | | |
| CC | (optimum 6-8), and retains more than 90% of its activity after 8 hours at | | |
| CC | 95 deg.C. The invention also provides gene sequences encoding a | | |
| CC | polypeptide of formula SIG-Ala-Gly-Ala-PHO, where SIG is a signal | | |
| CC | peptide from subtilisin, and PHO is the above protease. Host cells | | |
| CC | (especially <i>Bacillus</i> strains) transformed with vectors comprising the | | |
| CC | genes are used for the recombinant production of the protease. The | | |
| CC | hyperthermostable protease which can be prepared in quantity suitable for | | |
| CC | industrial use, can be used as an additive for drugs, washing agents and | | |
| CC | foodstuffs and for chemical synthesis | | |
| XX | Sequence 654 AA; | | |
| QO | | | |

| | |
|---|--------------------------------------|
| | 72.3% Score 2483.5 DB 2 Length 654 |
| Query Match | |
| Best Local Similarity | 72.1% Pred. No. 2.6e-144; |
| Matches 481; Conservative | 69; Mismatches 96; Indels 21; Gaps 8 |
| Oy | |
| 1 MKRIGAVVTLVLVGLAGTALAAPVKEV--VRNNAVOGQKNYGLITPGIPFKVKGRMNNO | 58 |
| : :: : : : | : : : : : : : : : : : : : : : : : : |

[illegible]

| | |
|-------------------------|--|
| LT 7 | |
| 4122 | |
| AAW24122 | standard; protein; 522 AA. |
| AAW24122; | |
| 17-OCT-2003 | (revised) |
| 20-APR-1998 | (first entry) |
| Pyrococcus furiosus | protease. |
| Protease; | research reagent; thermal stability; Pyrococcus furiosus |
| Pyrococcus furiosus; | DSM-3638. |
| Key | Location/Qualifiers |
| Misc-difference | 428 |
| /label= Gly, Val | |
| /note= "encoded by GNA" | |

PR 12-DEC-1995; 95JP-00323285.
 XX (TAKI) TAKARA SHUZO CO LTD.
 XX Takakura H, Morishita M, Yamamoto K, Mitra M, Aaeda K;
 XX Teunasaawa S, Kato I;
 XX WPI; 1997-332794/30.
 DR N-PSDB; AAT85666.
 XX
 XX Protease(s) and genes encoding them obtained from Thermococcus and
 PT Pyrococcus strains - have extremely high thermal stability and are useful
 PT industrially and as research reagents.
 XX
 PS Claim 5; Page 87-90; 159pp; Japanese.
 XX
 XX This sequence represents the protease from Pyrococcus furiosus DSM-3638.
 CC This sequence is a protease of the invention. The proteases of the
 CC invention have extremely high thermal stability. The proteases can be
 CC used as research reagents, and industrially in the food, drug and
 CC chemical industries. (Updated on 17-OCT-2003 to standardise OS field)
 CC
 XX
 SQ Sequence 522 AA;
 Query Match 62.2%; Score 2138.5; DB 2; Length 522;
 Best Local Similarity 78.3%; Pred. No. 3.3e-123;
 Matches 407; Conservative 36; Mismatches 70; Indels 7; Gaps 3;
 QY 141 SVSQIGADTVWNSLGYDGSQVVAIVDTGIDANHPDLKGVIGYDAVNGRSTPYDQGH 200
 DB 9 SAAQVMATYVMN-LGYDGSQGITIGIIDTGIDASHPDLQGVIGVDFVNGRSYPYDQGH 67
 QY 201 GTHVAGTGSV-NSQYGVAPGAKLVGKVLGADSGSVSTIIAGVDMVQNKXKG 259
 DB 68 GTHVASTIAAGGAASNGKYGMAFGAKLAGIKVLGADSGSISTIIKGEMAVNDKXKG 127
 QY 260 IRVNLISGSSQSDGDTLSQAVNNAMDAIGVCAVAGNSGPTTYVSGSPAASKVITY 319
 DB 128 IKVNLISGSSQSDGDTLSQAVNNAMDAIGVCAVAGNSGPTTYVSGSPAASKVITY 187
 QY 320 GAVDNDNIASFSSRGPTADGRLKEVVAAGVDIIAPRAGTSMGTPIINDYTKASGTS 379
 DB 188 GAVDKYDVTFSFSSRGPTADGRLKEVVAAGVDIIAPRAGTSMGTPIINDYTKASGTS 247
 QY 380 ATPHVSQVALLIOAHPSWTPDKYKTLIETADIVAPKETADIYAGRVVYKAIKXD 439
 DB 248 ATPHVAIGIALLLIOAHPSWTPDKYKTLIETADIVAPKETADIYAGRVVYKAIKXD 307
 QY 440 YAKLFTGSVADKGSATHTFDVSGATFVTATLYMDTSSDIDLTLYPDNGNEVDYSYTA 499
 DB 308 YAKLVFTGYVANKSQTHQFVIGASFTVATLYMDNANSDDLTLYPDNGNQVDYSYTA 367
 QY 500 YGFERYGYNFTAGTWTYKVVSYKCAANYQVDVYSDGSLSSGGGNPNPNPNPTPTD 559
 DB 368 YGFERYGYNFTAGTWTYKVVSYKCAANYQVDVYSDGSLSSGGGNPNPNPNPTPTD 422
 QY 560 TQTFGSDVNDWDTSDFTMAVNSGATKITGDLFTSYNDLTLYPDNGNLDVRSST 619
 DB 423 AKTFXSHYHYDRSDITMTVNSGATKITGDLFTSYNDLTLYPDNGNLDVRSST 482
 QY 620 NSYEHVEYANPAGTWTFLVVAISTYGNADYQKAVVYTG 659
 DB 483 NSYEHVEYANPAGTWTFLVVAISTYGNADYQKAVVYTG 522
 RESULT 8
 ID AAM94838
 XX AAM94838 standard; protein; 522 AA.
 AC AAM94838;
 XX
 XX 06-MAY-1999 (first entry)
 DT
 XX

DE Hyperthermostable protease.
 XX
 XX Hyperthermostable; protease; thermophilic; bacterium; subtilisin;
 KM additive; drug; washing agent; foodstuff; chemical synthesis.
 XX
 XX Pyrococcus furiosus.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 428
 FT /label= Gly or Val
 XX
 XX WO9856926-A1.
 XX
 XX 17-DEC-1998.
 XX
 XX 04-JUN-1998; 98WO-JP002465.
 XX
 XX 10-JUN-1997; 97JP-00151969.
 XX
 XX (TAKI) TAKARA SHUZO CO LTD.
 XX
 XX Takakura H, Morishita M, Shimojo T, Aaeda K, Kato I;
 XX WPI; 1999-080907/07.
 DR
 XX
 XX Recombinant hyperthermostable protease from Pyrococcus furiosus - and
 PT gene encoding it, for large scale production of the protease for
 PT industrial use.
 PT
 XX
 XX Claim 1; Page 39-41; 82pp; Japanese.
 PS
 XX
 XX The invention relates to a hyperthermostable protease derived from a
 CC thermophilic bacterium (especially Pyrococcus furiosus). The protease has
 CC working temperature 40-110 deg.C (optimum 80-95 deg.C), working pH 5-10
 CC (optimum 6-8), and retains more than 90% of its activity after 8 hours at
 CC 95 deg.C. The invention also provides gene sequences encoding a
 CC polypeptide of formula SIG-Ala-Gly-Gly-Ala-PRO, where SIG is a signal
 CC peptide from subtilisin, and PRO is the above protease. Host cells
 CC (especially Bacillus strains) transformed with vectors comprising the
 CC genes are used for the recombinant production of the protease. The
 CC hyperthermostable protease which can be prepared in quantity suitable for
 CC industrial use, can be used as an additive for drugs, washing agents and
 CC foodstuffs and for chemical synthesis
 CC
 XX
 SQ Sequence 522 AA;
 Query Match 62.2%; Score 2138.5; DB 2; Length 522;
 Best Local Similarity 78.3%; Pred. No. 3.3e-123;
 Matches 407; Conservative 36; Mismatches 70; Indels 7; Gaps 3;
 QY 141 SVSQIGADTVWNSLGYDGSQVVAIVDTGIDANHPDLKGVIGYDAVNGRSTPYDQGH 200
 DB 9 SAAQVMATYVMN-LGYDGSQGITIGIIDTGIDASHPDLQGVIGVDFVNGRSYPYDQGH 67
 QY 201 GTHVAGTGSV-NSQYGVAPGAKLVGKVLGADSGSVSTIIAGVDMVQNKXKG 259
 DB 68 GTHVASTIAAGGAASNGKYGMAFGAKLAGIKVLGADSGSISTIIKGEMAVNDKXKG 127
 QY 260 IRVNLISGSSQSDGDTLSQAVNNAMDAIGVCAVAGNSGPTTYVSGSPAASKVITY 319
 DB 128 IKVNLISGSSQSDGDTLSQAVNNAMDAIGVCAVAGNSGPTTYVSGSPAASKVITY 187
 QY 320 GAVDNDNIASFSSRGPTADGRLKEVVAAGVDIIAPRAGTSMGTPIINDYTKASGTS 379
 DB 188 GAVDKYDVTFSFSSRGPTADGRLKEVVAAGVDIIAPRAGTSMGTPIINDYTKASGTS 247
 QY 380 ATPHVSQVALLIOAHPSWTPDKYKTLIETADIVAPKETADIYAGRVVYKAIKXD 439
 DB 248 ATPHVAIGIALLLIOAHPSWTPDKYKTLIETADIVAPKETADIYAGRVVYKAIKXD 307
 QY 440 YAKLFTGSVADKGSATHTFDVSGATFVTATLYMDTSSDIDLTLYPDNGNEVDYSYTA 499
 DB 308 YAKLVFTGYVANKSQTHQFVIGASFTVATLYMDNANSDDLTLYPDNGNQVDYSYTA 367

QY 500 YGEKVGYNPTAGTWTWKVSYKGAANYQVDVSDGSLSSGGGNNPNPNPTPTD 559
 DB 368 YGEKVGYNPTDGTWTIKVSYSGSANYQVDVSDGSLSSG-----RGSSPSPPPTPTVD 422
 QY 560 TQFFTGSDVNDYMTSDTFPTMNVNSGATKTTGDLTFPTDSYNDLXYLDPNGLVDKSTSS 619
 DB 423 AKTFQXSDHYTYSRSDFTFTVNSGATKTTGDLVFTDTSYHDLXYLDPNQKLVDRSESP 482
 QY 620 NSYEHEVYANPAPGTWFLVYASTYGMADYOLKAVYYG 659
 DB 483 NSYEHEVYLTAPGTWFLVYAYTYGMYAYELTAKVYYG 522

RESULT 9

AAW94836
 ID AAW94836 standard; protein; 412 AA.

AAW94836;

DT 06-MAY-1999 (first entry)

DE Hyperthermostable protease fragment.

KM Hyperthermostable; protease; thermophilic; bacterium; subtilisin;
 additive; drug; washing agent; foodstuff; chemical synthesis.

OS Pyrococcus furiosus.

PN W09856926-A1.

PD 17-DEC-1998.

PF 04-JUN-1998; 98WO-JP002465.

PR 10-JUN-1997; 97JP-00151969.

PA (TAKI) TAKARA SHUZO CO LTD.

PI Takakura H, Morishita M, Shimajo T, Asada K, Kato I;

DR WPI; 1999-080907/07.

DR N-PSDB; AAX05920.

PT Recombinant hyperthermostable protease from *Pyrococcus furiosus* - and
 gene encoding it, for large scale production of the protease for
 industrial use.

PS Claim 2; Page 35-37; 82pp; Japanese.

CC The invention relates to a hyperthermostable protease derived from a
 thermophilic bacterium (especially *Pyrococcus furiosus*). The protease has
 working temperature 40-110 deg.C (optimum 80-95 deg.C), working pH 5-10
 (optimum 6-8), and retains more than 90% of its activity after 8 hours at
 95 deg.C. The invention also provides gene sequences encoding a
 polypeptide of formula $\text{SIG-Ala-Gly-Gly-Asn-PRO}$, where SIG is a signal
 peptide from subtilisin, and PRO is the above protease. Host cells
 (especially *Bacillus* strains) transformed with vectors comprising the
 genes are used for the recombinant production of the protease. The
 hyperthermostable protease which can be prepared in quantity suitable for
 industrial use, can be used as an additive for drugs, washing agents and
 foodstuffs and for chemical synthesis

XX Sequence 412 AA;

QY Query Match 49.7%; Score 1707; DB 2; Length 412;

Best Local Similarity 80.6%; Pred. No. 8,7e-97;
 Matches 325; Conservativity 29; Mismatches 47; Indels 2; Gaps 2;

QY 141 SVSGIGADTYWNSLIGYSGGVVAIVTGTIDANHPDLKGYIGYADVANGRSTPYDQGH 200
 DB 9 SAQGVNATYYWN-LGYDGSIGITIGITIDTIDASHPDQKVTIGVNDVNGRSYFYDDGH 67

QY 201 GTHVAGIVAGTGV-NSQYIYGAAPGAKLVGVKVLGADGGSGSVSTIAGVDMVYONKDKG 259
 DB 68 GTHVASIAAGTGAASNGKYGMAPGAKLAGIKVLGADGGSGSISTIIKGVEMAVDNDKYG 127
 QY 260 IRYINISLSSGSSDSDTSLSSQAVNNAMPAGIVVCVAAAGSGPNTYVGS PAASKVITY 319
 DB 128 IKVINISLSSGSSDSDTSLSSQAVNNAMPAGIVVCVAAAGSGPNTYVGS PAASKVITY 187
 QY 320 GAVDSDNINIAFSRSPRTADGRLEPVPVAGVDIIAPRASGTSMPINDYTTKASGTSM 379
 DB 188 GAVDKDVTITSPSSRPTADGRLEPVPVAGVDIIAPRASGTSMPINDYTTKASGTSM 247
 QY 380 ATPHVSQVALLIQAPSWTPDKVKTALLETADIVAPKEIADIAYAGRVNVAIKAYDD 439
 DB 248 ATPHVAGIALLLQAPSWTPDKVKTALLETADIVAPKEIADIAYAGRVNVAIKAYDD 307
 QY 440 YAKLTFTGSVADGSAITHTPDVSGATFYVATLYMDGSSPIDLYDPNENEDYSTAY 499
 DB 308 YAKLTFTGVVANGSQTHQFVIGSASFVATLYMDANSPLDLYDPNENEDYSTAY 367
 QY 500 YGEKVGYNPTAGTWTWKVSYKGAANYQVDVSDGSLSSG 542
 DB 368 YGEKVGYNPTDGTWTIKVSYSGSANYQVDVSDGSLSSG 410

RESULT 10

AAW87009
 ID AAW87009 standard; protein; 237 AA.

AAW87009;

DT 10-MAY-1996 (first entry)

DE Hyperthermostable protease.

KM Protease; hyperthermostable; thermostability.

OS Pyrococcus furiosus.

PN W09534645-A1.

PD 21-DEC-1995.

PF 05-JUN-1995; 95WO-JP001095.

PR 13-JUN-1994; 94JP-00130236.

PR 26-JUL-1994; 94JP-00173912.

PA (TAKI) TAKARA SHUZO CO LTD.

PI Mita M, Yamamoto K, Morishita M, Asada K, Teunasa S, Kato I;

DR WPI; 1996-049674/05.

DR N-PSDB; AAT08131.

XX Pyrococcus furiosus hyperthermostable protease gene - useful for
 recombinant prodn. of hyperthermostable protease.

XX Example 3; Page 70-71; 85pp; Japanese.

CC The invention relates to the hyperthermostable protease of *Pyrococcus*
 furiosus and its prodn. as a recombinant protein in transformants using a
 vector carrying the protease gene (AAT08131). A genomic DNA sequence of
 the invention is given in AAT08131 and its encoded protein in AAW87009

XX Sequence 237 AA;

QY Query Match 33.6%; Score 1154; DB 2; Length 237;

Best Local Similarity 97.4%; Pred. No. 4,7e-63;

QY 9 SAQGVNATYYWN-LGYDGSIGITIGITIDTIDASHPDQKVTIGVNDVNGRSYFYDDGH 67

Matches 224; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 176 DLKGVIGWYDAVNRSTPYDDQGHGTIVAGTGSVNSQYIGVAPGAKLVGVKVLGA 235
 DB 1 DLKGVIGWYDAVNRSTPYDDQGHGTIVAGTGSVNSQYIGVAPGAKLVGVKVLGA 60

QY 236 DGSVSTIITAGVWVYVQNKDKYIRVYNLSIGSSQSDGTDSLQAVNNAMDAGIVVCV 295
 DB 61 DGSVSTIITAGVWVYVQNKDKYIRVYNLSIGSSQSDGTDSLQAVNNAMDAGIVVCV 120

QY 296 AAGNGPMTYTVGSPAAASKVITVGAVDNDNIASFSSRGPTADRLKPEVVAEGVDIIA 355
 DB 121 AAGNGPMTYTVGSPAAASKVITVGAVDNDNIASFSSRGPTADRLKPEVVAEGVDIIA 180

QY 356 PRASGTSMTPIINDYTTTASGTSMTAPVSGVGLIIQAHPSMTPEDKYT 405
 DB 181 PRASGTSMTPIINDYTTTASGTSMTAPVSGVGLIIQAHPSMTPEDKYT 230

RESULT 11

AAR87011
 ID AAR87011 standard; peptide; 188 AA.

AC AAR87011;

DT 10-MAY-1996 (first entry)

DE Peptide sequence.

DE Protease; hyperthermostable; thermostability.

OS Pyrococcus furiosus.

XX WO9534645-A1.

XX 21-DEC-1995.

XX 05-JUN-1995; 95MO-JP001095.

XX 13-JUN-1994; 94JP-00130236.

XX 26-JUL-1994; 94JP-00173912.

XX (TAKI) TAKARA SHUZO CO LTD.

PI Mitra M, Yamamoto K, Morishita M, Asada K, Terasawa S, Kato I;

XX WPI; 1996-049674/05.

XX Pyrococcus furiosus hyperthermostable protease gene - useful for

XX recombinant prodn. of hyperthermostable protease.

XX Example 4; Page 66-67; 85pp; Japanese.

XX The invention relates to a gene (AAT08141) that codes for a

XX hyperthermostable protease (AAR87007) of Pyrococcus furiosus. 2 DNA

XX sequences (AAT08133-34) are provided encoding peptides (AAR87010-11)

XX Sequence 188 AA;

Query Match 27.2%; Score 935; DB 2; Length 188;

Best Local Similarity 98.4%; Pred. No. 1e-49;

Matches 184; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 200 HGTIVAGTGSVNSQYIGVAPGAKLVGVKVLGADSGSVSTIIAGVWVYVQNKDKY 259
 DB 2 HGTIVAGTGSVNSQYIGVAPGAKLVGVKVLGADSGSVSTIIAGVWVYVQNKDKY 61

QY 260 IRVINISLSSQSDGADSLQAVNNAMDAGIVVCVAAAGNSGPTTYTGSPPAAASKVITV 319
 DB 62 IRVINISLSSQSDGADSLQAVNNAMDAGIVVCVAAAGNSGPTTYTGSPPAAASKVITV 121

QY 320 GAVDSNDNIASFSSRGPTADRLKPEVVAEGVDIIAPRASGTSMTPIINDYTTKASGTS 379
 DB 182 GAVDSNDNIASFSSRGPTADRLKPEVVAEGVDIIAPRASGTSMTPIINDYTTKASGTS 379

DB 122 GAVDSNDNIASFSSRGPTADRLKPEVVAEGVDIIAPRASGTSMTPIINDYTTKASGTS 181
 QY 380 ATPHVS 386
 DB 182 ATPHVS 188

RESULT 12

AAM24126
 ID AAM24126 standard; protein; 188 AA.

AC AAM24126;

DT 17-OCT-2003 (revised)

DT 20-APR-1998 (first entry)

DE Thermococcus protease fragment.

XX Protease; research reagent; thermal stability; thermococcus celer.

OS Thermococcus celer; DSM-2476.

XX WO9721823-A1.

XX 19-JUN-1997.

XX 07-NOV-1996; 96MO-JP003253.

XX 12-DEC-1995; 95JP-00323285.

XX (TAKI) TAKARA SHUZO CO LTD.

PI Takakura H, Morishita M, Yamamoto K, Mitra M, Asada K;

XX Terasawa S, Kato I;

XX WPI; 1997-332794/30.

XX N-PSDB; AAT85676.

XX Protease(s) and genes encoding them obtained from Thermococcus and

XX Pyrococcus strains - have extremely high thermal stability and are useful

XX industrially and as research reagents.

XX Disclosure; Page 110-112; 159pp; Japanese.

XX This sequence represents a fragment of the protease from Thermococcus

XX celer DSM-2476 (see AAM24121 for full length sequence). This sequence is

XX a fragment of the protease of the invention. The proteases of the

XX invention have extremely high thermal stability. The proteases can be

XX used as research reagents, and industrially in the food, drug and

XX chemical industries. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 188 AA;

Query Match 27.2%; Score 935; DB 2; Length 188;

Best Local Similarity 98.4%; Pred. No. 1e-49;

Matches 184; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 200 HGTIVAGTGSVNSQYIGVAPGAKLVGVKVLGADSGSVSTIIAGVWVYVQNKDKY 259
 DB 2 HGTIVAGTGSVNSQYIGVAPGAKLVGVKVLGADSGSVSTIIAGVWVYVQNKDKY 61

QY 260 IRVINISLSSQSDGADSLQAVNNAMDAGIVVCVAAAGNSGPTTYTGSPPAAASKVITV 319
 DB 62 IRVINISLSSQSDGADSLQAVNNAMDAGIVVCVAAAGNSGPTTYTGSPPAAASKVITV 121

QY 380 ATPHVS 386
 DB 182 ATPHVS 188

[illegible]

| | | | | |
|-----------|--|---|---|-------|
| Oy | 449 | VAD----- | -----KGSATHTFEDVSGATF-VTATY---WPTGSSDIDLVLVDP | 487 |
| | | | : : : : | : : |
| Db | 398 | NIDVPMHTYISGLYPGSRKYSDDTWTFMNTSTSPYIAITLITLIDPMWANYNFPDIYLYDP | | 454 |
| RESULT 14 | | | | |
| XX | AAW13667 | | | |
| XX | AAW13667 | standard; protein, 734 AA. | | |
| XX | | | | |
| XX | AAW13667; | | | |
| XX | | | | |
| DT | 06-OCT-1997 | (first entry) | | |
| XX | | | | |
| DE | Streptomyces viridosporus dhpa gene product. | | | |
| XX | | | | |
| XX | asymmetric hydrolase, dhpa; 4-substituted-1,4-dihydropyridine; | | | |
| KW | derivative; Streptomyces viridosporus; ester; chiral; synthesis; | | | |
| KW | cardiovascular; treatment; hypertension; ischaemic heart disease. | | | |
| XX | | | | |
| OS | Streptomyces viridosporus. | | | |
| XX | | | | |
| FH | Key | Location/Qualifiers | | |
| FT | Peptide | 205..724 | | |
| XX | | /note="fragment of dhpa protein; see AAW13666" | | |
| XX | WO9705243-A1. | | | |
| XX | | | | |
| PD | 13-FEB-1997. | | | |
| XX | | | | |
| XX | 30-JUL-1996; | 96WO-JP002147. | | |
| PF | | | | |
| PR | 31-JUL-1995; | 95JP-00212975. | | |
| XX | | | | |
| PR | 29-FEB-1996; | 96JP-00067478. | | |
| XX | | | | |
| PA | (SAOC) MERCIAN CORP. | | | |
| XX | | | | |
| PI | Arisawa A, Matsufuji M, Tsuruta T, Dobashi K, Nakashima T; | | | |
| PI | Ieshiki K, Yoshioka T; | | | |
| XX | | | | |
| DR | WPI; 1997-145682/13. | | | |
| XX | | | | |
| DR | N-PSDB; AAT61454. | | | |
| PT | Asymmetric hydrolase gene derived from Streptomyces viridosporus - acts | | | |
| PT | on 4-substituted-1,4-di:hydro:pyridine derivatives to produce chiral | | | |
| PT | derivatives useful for synthesis of cardiovascular drugs. | | | |
| XX | | | | |
| XX | Claim 3; Page 49-55; 78pp; Japanese. | | | |
| XX | | | | |
| CC | This sequence is an asymmetric hydrolase encoded by the Streptomyces | | | |
| CC | viridosporus dhpa gene. The enzyme acts on 4-substituted-1,4- | | | |
| CC | dihydropyridine derivatives. The enzyme allows the efficient conversion | | | |
| CC | of 4-substituted-1,4-dihydropyridine esters to chiral partially | | | |
| CC | hydrolysed derivatives, for use in the synthesis of cardiovascular drugs | | | |
| CC | suitable for the treatment of e.g. hypertension and ischaemic heart | | | |
| CC | disease | | | |
| XX | | | | |
| XX | Sequence 734 AA; | | | |
| SO | | | | |
| Oy | Query Match | 19.3%; Score 665; DB 2; Length 734; | | |
| | Best Local Similarity | 31.2%; Pred. No. 2.3e-32; | | |
| Db | Matches 216; Conservative % 84; Mismatches 269; Indels 124; Gaps 26; | | | |
| Oy | 22 | LAARVAVVRNNAVQOKNYGLTGLPKKVOORMNMNEVDVIMFGSYGDDRAVKVURL | 81 | |
| | | : : : : | : : | |
| Db | 97 | LPAPAAALVAVSGKDRRLPDTTEIG--KAATRNQKQGLKVIV--GTQGAAPAAKEVR | 151 | |
| Oy | 82 | MGAOVKRSYKIIIPAAVKIKARLDLLIAGMIDTGYFGATRVSGIKFTQEDYKVOVDAT5 | 141 | |
| | | : : : : | : : | |
| Db | 152 | EAGLRLRTIINSUNDVAVRTPHED---ASELMDVATVNDGRTRTSGIAHWLDDVTRAAALDT5 | 208 | |
| Oy | 142 | VSQIGADTVNNSLGIDGSGVVAIVDTGIDANHPDLKSGVIGWYDAVNGRSTP--YDDQG | 199 | |
| | | : : : : | : : | |

```

Db      209 VQGIAPKPKAM-SAGYDGKGVKIAVLDTGVDTSHDPLKGRVTA---SKNFTAAPGAGDKVG 264
Qy      200 HGTHTVAGIAGTGSVN-SQYIGVAPGAKLVGVKVLGADSGSVSTIIAGVMPVNVQNKDKY 258
Db      265 HGTHTVASTAAGTGAQSGKGYKGVAPGAALINKKVLDDSGFQDSDGILLGMEHMAAQ---- 320
Qy      259 GIRVYNLSIGSSQSDGTDLSQAVNN-AMDAGIVVCYAAGNSGPNYTVGSPAAASKVI 317
Db      321 GADVNMISIGGMDTPE-TDPLEAAVDKLSAEKGVFAIAAGNEGPE--SIGSPGSADAL 377
Qy      318 TVGAVDSNDNINASFSSRGP-TADGRLKPEVVAPOVDIIAPRASGTSMTGPIND--YTTK 373
Db      378 TVGAVDDKDKLADFSSTGPRLDGAIKPDVTAPOVDITPAASAGNDIQEVEGEPAGYMT 437
Qy      374 ASGTSMAIPHVSGVALLIOAHPSWTPDKVKIALIETADIAPKEIADIAGAGRVNYYK 433
Db      438 ISGTSMAIPHVGAALALKQCHPDMWTSALIKGAL--TGSTGKGK-YTFPBGSSGRIQDK 494
Qy      434 AIK-----YDYAKLFTTGSVADKGSATH 457
Db      495 ALQQTVIADPVSVSPGVQOWPHTDDEPYTKOLTYRNLTGTOVTLKLTSTADPKKAAPA 554
Qy      458 TFDVSGATFVTATLYMDTGSSDIDLXYLDPNGNEVDYSYTAAYGEEKGYNNPTAGTWT 517
Db      555 GFPTLGATTTVPA---GGSASVDMTADTRLGGTVDGAYSAYVAATGGQTVRTAAVQR 611
Qy      518 KVVSY-----KGAANYQVDVSDGSLSGSGGPNPNPNPTPTDTQT--- 563
Db      612 EVESYDVTVRHIGRDKPTTEHLTDLIGVAGLSGRGYG-----APATDTATLRLP 662
Qy      564 TGS-VNDVMDTSDFPTM-----NVNSGATKXTGDLTFPT-SYNDLDTLYLDPNGNLV 613
Db      663 KGYTLVDSWIAKDQETLKGIDMLVQPLSVTKOT-TLTLDARTTKADITVDPK----- 717
Qy      614 DRSTSSNVEHEVYANPAPGTWTFVLVAYSTYG 646
Db      718 -----AKPLSAT---IGTYTDTAG 733

RESULT 15
AA013668
ID      AA013668 standard; protein; 823 AA.
XX
AC      AA013668;
XX
DT      17-OCT-2003 (reviewed)
DT      06-OCT-1997 (first entry)
XX
DB      DhpA-mel chimeric protein.
XX
KW      asymmetric hydrolase; dhpA; 4-substituted-1,4-dihydropyridine;
KW      derivative; Streptomyces viridoporus; ester; chiral; synthesis;
KW      cardiovascular; treatment; hypertension; ischaemic heart disease.
XX
OS      Streptomyces viridoporus.
OS      Streptomyces antibioticus.
OS      Chimeric.
XX
FH      Key
FH      Protein
FT      /label= dhpA_protein_product
FT      /label= 724
FT      /note= "see AA013666"
FT      /note= 735..823
FT      /note= "melanin (partial sequence)"
XX
XX      MO9705243-A1.
XX      13-FEB-1997.
XX      30-JUL-1996; 96MO-JP002147.
XX      31-JUL-1995; 95JP-00212975.

```

```

PR      29-FEB-1996; 96JP-00067478.
XX
XX      (SAOC ) MERCIAN CORP.
XX
PI      Aritaawa A, Matsufuji M, Tsuruta T, Dobashi K, Nakashima T,
PI      Ieshiki K, Yoshiooka T;
XX
XX      WPI: 1997-145682/13.
XX      N-PSDB; AAT61455.
XX
FT      Asymmetric hydrolase gene derived from Streptomyces viridoporus - acts
FT      on 4-substituted-1,4-dihydropyridine derivatives to produce chiral
PT      derivatives useful for synthesis of cardiovascular drugs.
XX
PS      Claim 5; Page 37-43; 78pp; Japanese.
XX
CC      This sequence is a fusion protein comprising Streptomyces viridoporus
CC      dhpA gene product (an asymmetric hydrolase) which acts on 4-substituted-
CC      1,4-dihydropyridine derivatives, and melanin from S. antibioticus. The
CC      dhpA enzyme allows the efficient conversion of 4-substituted-1,4-
CC      dihydropyridine esters to chiral partially hydrolysed derivatives, for
CC      use in the synthesis of cardiovascular drugs suitable for the treatment
CC      of e.g. hypertension and ischaemic heart disease. (Updated on 17-OCT-2003
CC      to standardise OS field)
XX
SQ      Sequence 823 AA;

```

```

Query Match      19.3%; Score 665; DB 2; Length 823;
Best Local Similarity 31.2%; Pred. No. 2, 6e-32;
Matches 216; Conservative 84; Mismatches 269; Indels 124; Gaps 26;

```

```

Qy      22 LAAPVPRVNNANVOOKNYGLTPGLFKYKORNNNOEVDYIMFGSYGDRRAVNVRL 81
Db      97 LPADARIVASGLDRLDITELG--KAATNSQSGKAVIV--GYGAARAAKAEVR 151
Qy      82 MGAQVYKSYKILPAVAVKIKADLLIAGMTDTGYFGNTRVSGIKETQEDYKQVDDATG 141
Db      152 EAGELRRTLTSINADAVRTPHED--ASELMADVNTGDRTAGIAHVMLDGVRRAALDTS 208
Qy      142 VSGIGADVTWNSLGYVSGVNVVAIVDTGIDAHNPLDKGVIMYDAVNRSTP--YDDOG 199
Db      209 VQGIAPKPKAM-SAGYDGKGVKIAVLDTGVDTSHDPLKGRVTA---SKNFTAAPGAGDKVG 264
Qy      200 HGTHTVAGIAGTGSVN-SQYIGVAPGAKLVGVKVLGADSGSVSTIIAGVMPVNVQNKDKY 258
Db      265 HGTHTVASTAAGTGAQSGKGYKGVAPGAALINKKVLDDSGFQDSDGILLGMEHMAAQ---- 320
Qy      259 GIRVYNLSIGSSQSDGTDLSQAVNN-AMDAGIVVCYAAGNSGPNYTVGSPAAASKVI 317
Db      321 GADVNMISIGGMDTPE-TDPLEAAVDKLSAEKGVFAIAAGNEGPE--SIGSPGSADAL 377
Qy      318 TVGAVDSNDNINASFSSRGP-TADGRLKPEVVAPOVDIIAPRASGTSMTGPIND--YTTK 373
Db      378 TVGAVDDKDKLADFSSTGPRLDGAIKPDVTAPOVDITPAASAGNDIQEVEGEPAGYMT 437
Qy      374 ASGTSMAIPHVSGVALLIOAHPSWTPDKVKIALIETADIAPKEIADIAGAGRVNYYK 433
Db      438 ISGTSMAIPHVGAALALKQCHPDMWTSALIKGAL--TGSTGKGK-YTFPBGSSGRIQDK 494
Qy      434 AIK-----YDYAKLFTTGSVADKGSATH 457
Db      495 ALQQTVIADPVSVSPGVQOWPHTDDEPYTKOLTYRNLTGTOVTLKLTSTADPKKAAPA 554
Qy      458 TFDVSGATFVTATLYMDTGSSDIDLXYLDPNGNEVDYSYTAAYGEEKGYNNPTAGTWT 517
Db      555 GFPTLGATTTVPA---GGSASVDMTADTRLGGTVDGAYSAYVAATGGQTVRTAAVQR 611
Qy      518 KVVSY-----KGAANYQVDVSDGSLSGSGGPNPNPNPTPTDTQT--- 563
Db      612 EVESYDVTVRHIGRDKPTTEHLTDLIGVAGLSGRGYG-----APATDTATLRLP 662
Qy      564 TGS-VNDVMDTSDFPTM-----NVNSGATKXTGDLTFPT-SYNDLDTLYLDPNGNLV 613

```

Db 663 KGTYLVDSWIANDFGTLKGGIDWLVOFKLSVTKDT-TLTLIDARTTKADITVPDPK----- 717
Qy 614 DRSTSSNSYEHVEYANPAPGWTPELVYASTYG 646
Db 718 -----AKPLSAT---IGTYDTAG 733

Search completed: January 6, 2007, 22:24:53
Job time : 144 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2007 Bioceleration Ltd.

OW protein - protein search, using sw model

Run on: January 6, 2007, 22:25:12 ; Search time 27 Seconds

(without alignments)
2348.402 Million cell updates/sec

Title: US-10-800-684-1

Perfect score: 3437

Sequence: 1 MKRLGAVVLAIVLVLGAGT.....YAVSTGYMADYQLKAVVYGG 659

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: pirl:*
2: pirl2:*
3: pirl3:*
4: pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 669.5 | 19.5 | 442 | 2 A69587 | intracellular alka |
| 2 | 637 | 18.5 | 444 | 2 B83891 | intracellular alka |
| 3 | 557.5 | 16.2 | 379 | 1 SUBSCL | subtilisin (EC 3.4 |
| 4 | 547 | 15.9 | 382 | 1 I39780 | subtilisin (EC 3.4 |
| 5 | 533.5 | 15.5 | 382 | 1 SUBSN | subtilisin (EC 3.4 |
| 6 | 533.5 | 15.5 | 1398 | 2 T28159 | pyrolysin (EC 3.4 |
| 7 | 524.5 | 15.3 | 274 | 1 SUBSD | subtilisin (EC 3.4 |
| 8 | 522.5 | 15.2 | 806 | 2 A41341 | microbial serine p |
| 9 | 519.5 | 15.1 | 275 | 2 JC1085 | subtilisin (EC 3.4 |
| 10 | 515.5 | 15.0 | 384 | 2 JCA802 | alkaline proteinase |
| 11 | 513.5 | 14.9 | 381 | 2 JH0778 | subtilisin (EC 3.4 |
| 12 | 512.5 | 14.9 | 381 | 1 SUBSI | subtilisin (EC 3.4 |
| 13 | 510.5 | 14.9 | 380 | 2 A49778 | high-alkaline seri |
| 14 | 510 | 14.8 | 1433 | 1 A36734 | backloopeptidase F |
| 15 | 509.5 | 14.8 | 381 | 1 SUBSS | subtilisin (EC 3.4 |
| 16 | 509.5 | 14.8 | 381 | 2 JQ1487 | subtilisin (EC 3.4 |
| 17 | 497.5 | 14.5 | 272 | 2 A23624 | subtilisin (EC 3.4 |
| 18 | 487 | 14.5 | 757 | 2 C84120 | subtilisin-type pr |
| 19 | 495 | 14.4 | 374 | 2 I39781 | subtilisin (EC 3.4 |
| 20 | 491.5 | 14.3 | 378 | 2 A33973 | high-alkaline seri |
| 21 | 488.5 | 14.2 | 279 | 1 SUMYTV | thermase (EC 3.4 |
| 22 | 482 | 14.0 | 279 | 2 I39974 | serine proteinase |
| 23 | 477.5 | 13.9 | 799 | 2 G83753 | subtilisin-type pr |
| 24 | 473.5 | 13.8 | 535 | 2 B82358 | alkaline serine pr |
| 25 | 473.5 | 13.8 | 715 | 2 JCA408 | alkaline serine pr |
| 26 | 472 | 13.7 | 513 | 1 A35742 | subtilisin (EC 3.4 |
| 27 | 470 | 13.7 | 361 | 2 A48373 | high-alkaline seri |
| 28 | 470 | 13.7 | 361 | 2 G83756 | subtilisin-type al |
| 29 | 464.5 | 13.5 | 534 | 1 JS0173 | alkaline proteinase |

| | | | | | |
|----|-------|------|-----|----------|---------------------|
| 30 | 461 | 13.4 | 627 | 2 D75393 | serine proteinase, |
| 31 | 452 | 13.2 | 440 | 2 H72784 | probable alkaline |
| 32 | 445.5 | 13.0 | 519 | 2 S71451 | halolysin R4 (EC 3 |
| 33 | 443 | 12.9 | 488 | 2 A11930 | proteinase (import |
| 34 | 437.5 | 12.7 | 372 | 2 D83735 | subtilisin-type al |
| 35 | 437.5 | 12.7 | 645 | 1 SUBSMP | serine proteinase |
| 36 | 432.5 | 12.6 | 397 | 2 JMW075 | cysteine-dependent |
| 37 | 420.5 | 12.2 | 588 | 2 C83836 | subtilisin-type pr |
| 38 | 418.5 | 12.2 | 401 | 2 A57690 | aerolysin precursor |
| 39 | 417 | 12.1 | 747 | 2 T06580 | subtilisin-like pr |
| 40 | 412.5 | 12.0 | 467 | 1 S45493 | serine proteinase |
| 41 | 407.5 | 11.9 | 745 | 2 UC6119 | subtilisin-like pr |
| 42 | 399.5 | 11.6 | 326 | 1 C41335 | microbial serine p |
| 43 | 399 | 11.6 | 319 | 2 I39866 | alkaline serine p |
| 44 | 398.5 | 11.6 | 525 | 2 G84406 | halolysin (importe |
| 45 | 398 | 11.6 | 754 | 2 T06579 | subtilisin-like pr |

ALIGNMENTS

RESULT 1

A69587 intracellular alkaline serine proteinase aprx - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #ext_change 09-Jul-2004

C/Accession: A69587

R/Kmet, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertier, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, A.; Enllich, S.D.; Emerson, P.T.; Enllich, K.D.; Errington, J.; Ferrar, C.; Ferrar, E. Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapins, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Maue, Y. M.; Ogawa, K.; Ogawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanton, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, A.; Authors: Tanakoshi, A.; Tanaka, T.; Tetsura, P.; Tognoli, A.; Tosato, V.; Uchiyama, K.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K. A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A. A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis. A/Reference number: A69580; PMID:98044033; PMID:9384377

A/Accession: A69587

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-442 <KUN>

A/Cross-references: UNIPROT:O31788; UNIPARC:UPI00006043F; GB:299113; GB:AL009126; NID:G

A/Experimental source: strain 168

C/Genetics:

A/Gene: aprX

F/146-398/Domain: subtilisin homology <SBT>

Query Match 19.5%; Score 669.5; DB 2; Length 442;

Best local similarity 46.0%; Pred. No. 2.2e-29;

Matches 151; Conservative 51; Mismatches 103; Indels 23; Gaps 9;

| | | | |
|----|-----|---|-----|
| QY | 127 | FIQEDYKQVDDATVSQIGADTVNWSLGYDSSGVVAIVDTGIDANHPDLKQVIGWD | 186 |
| DB | 116 | YINREVKALDPTATEASH-AKEVVRNGQTLTKGVAVAVDGI-YPHDLERIRIIGPAD | 173 |
| QY | 187 | ANGASTPYDDGCHTHVAGIAGGSVNS-QYIGVAPAKVGVGVAGDSSGSTII | 245 |
| DB | 174 | MNOKTEPYDDGCHTHVAGIAGGSVNS-QYIGVAPAKVGVGVAGDSSGSTII | 245 |
| QY | 246 | AGVDWVQ-NKDK--YGIKVINLSIGSS--OSSDGTSLQAVVNNMAGIIVCVAAGN | 299 |
| DB | 234 | EGVENCIQYNEONPDEPIDIMSGDGLRDRHEDPLVAVAEAMAGIIVCVAAGN | 293 |
| QY | 300 | SGBNYTVVSGPAASRVITVGAVDN-----DNIAFSGKPTADGLKBEVVAQVD | 352 |
| DB | 294 | SGDSDQTIASPGVSEKVIIVGALDDNMTASDDDTVASPSSGPTVYGEKEDIIAPGVN | 353 |

C:Keywords: hydrolase; serine proteinase
F:136-342/Domain: subtilisin homology <SBT>

Query Match 15.9%; Score 547; DB 2; Length 382;

Best Local Similarity 34.5%; Pred. No. 8.8e-23; Mismatches 145; Indels 74; Gaps 17;

Matches 154; Conservative 73; Mismatches 145; Indels 74; Gaps 17;

QY 1 MKRLGAVV-ALVVLGLAGTALAPKPVVNNAAVQKXVGLTTPGLPKVQRMNN 57
DB 1 MKKLFTKVASAALLISLTATSVSAE-----EQKKQYLI--GFENQLQVTERV 48
QY 58 QEVD---TVIMFGSYGRDRAVKVLRMLNGAOKYSYKIIIPAVAVKIKARDLLIAGMTD 113
DB 49 EESDKQSGSMSPFAEVNDESIEMELL-----YEFEDIPVVSVELSPEDVDL----- 95
QY 114 TGYFGNTRVSGIKFIQEDYKNOVDAT---VSQIGADTVNMSLGDGSGVVAIVDTGI 170
DB 96 -----EKDPSTIYEDIEVITITNQVTWGITRVAQAPFAM--TRGYTGVRVAVLDTSI 148
QY 171 DANHPDLKGVIGWVDVANGRSTPYVDQGHGTHVAGTVAGTGSVNSQYGVAPGAKLVGV 230
DB 149 -STHPDL--NIRGCVSFVGPSPSYDQNGHGHVAGTIALNN-SIGVGVAPNAELIYV 204
QY 231 KYLGADSGSVSTIAGVDMVYQNKXGIRVNLSTGSSQSDGTDLSQAANNAMDAG 290
DB 205 KYLGANGSGSVSIAQGLQMTAQN---NIHVANLSTGSPV---GSQTELVAVNQATNAG 257
QY 291 IVYCVAAAGSGPNTYVSGPAAASKVITYGAVDSNDNIASFSSRGPETADGRLKEPVVAG 350
DB 258 VLVVAVATGNGSG--TVSPARYANALAVGATDQNNRASFQOYGTGLN-----IVAG 309
QY 351 VIIIPRAGTSMGPIFYNYTKASGTSMATPHVSGVGLIIQAPHSWTPDKYKALLET 410
DB 310 VGI-----QSTYPGNRVYASLSTSTNATPHVAGVAAALVKOKNPSMNTQIRHLIST 360
QY 411 ADIVAPKEIADIYAGRVNVYKAIR 436
DB 361 ATSLGNSN---QFGSGLVNAAEATR 382

RESULT 5
SUBSIN (EC 3.4.21.62) BPN, precursor - Bacillus amyloliquefaciens
N:Alternates names: subtilisin Novo
C:Species: Bacillus amyloliquefaciens
C:Date: 24-Apr-1984 #sequence revision 28-Aug-1985 #text change 05-Oct-2004
C:Accession: B25415; A93495; T44584; A92033; A00970
R:Vaanantha, N.; Thompson, L.D.; Rhodes, C.; Banner, C.; Nagle, J.; Filipula, D.
J. Bacteriol. 159, 811-819, 1984
A:Title: Genes for alkaline protease and neutral protease from Bacillus amyloliquefaciens
A:Reference number: A25415; MUID:85006739; PMID:6090391
A:Accession: B25415
A:Molecule type: DNA
A:Residues: 1-382 <VAS>
A:Cross-references: UNIPROT:P00782; UNIPARC:UPI0000136180; GB:K02496; NID:g142525; PIDN:
A:Experimental source: ATCC 23844
R:Wells, J.A.; Ferrari, E.; Henner, D.J.; Estell, D.A.; Chen, E.Y.
Nucleic Acids Res. 11, 7911-7925, 1983
A:Title: Cloning, sequencing, and secretion of Bacillus amyloliquefaciens subtilisin in
A:Reference number: A93495; MUID:84069812; PMID:6316278
A:Accession: A93495
A:Molecule type: DNA
A:Residues: 1-382 <WE2>
A:Cross-references: UNIPARC:UPI0000136180; EMBL:X00165; NID:g33337; PIDN:CAA24990.1; PID
R:Marland, F.S.; Smith, E.L.
J. Biol. Chem. 242, 5198-5211, 1967
A:Title: Subtilisin BPN'. VII. Isolation of cyanogen bromide peptides and the complete A
A:Reference number: A92033; MUID:68086682; PMID:6065094

A:Accession: A92033
A:Molecule type: protein
A:Residues: 108-162, 'PN', 165-167, 'D', 169-194, 'SA', 197-204, 'DA', 207-264, 'ST', 267-357, 'Q', 3
A:Cross-references: UNIPARC:UPI000011292B
R:Kratz, J.
In The Enzymes, 3rd ed., vol. 3, Boyer, P.D., ed., pp. 547-560, Academic Press, New York, 1
A:Title: Subtilisin: X-ray structure.
A:Reference number: A94443
A:Contents: annotation; X-ray crystallography, 2.5 angstroms; active site
A:Comment: Secretion of subtilisin is associated with the onset of sporulation, and many
not necessary for normal sporulation.

C:Genetics:
A:Start codon: GTG
C:Superfamily: Subtilisin; subtilisin homology
C:Keywords: hydrolase; serine proteinase
F:1-37/Domain: signal sequence #status predicted <SIG>
F:133-107/Domain: activation peptide #status predicted <APT>
F:108-382/Product: subtilisin BPN' #status experimental <MPT>
F:130-342/Domain: subtilisin homology <SBT>
F:139,171,388/Active site: Asp, His, Ser #status experimental

Query Match 15.5%; Score 533.5; DB 1; Length 382;
Best Local Similarity 38.5%; Pred. No. 4.8e-22;
Matches 143; Conservative 52; Mismatches 113; Indels 63; Gaps 14;

QY 78 VRLMGAQVQKYSYKIIIPV-----AVKIKARDLLIAGMTDITGYFGNTRVSGIKFIQ 130
DB 59 VISEKGVQKQKPYVDAAATLNEKAYELKD-----PSVAYVE 100
QY 131 DYKQVDDAT---SVSQIGADTVNMSLGDGSGVVAIVDTGIDANHPDLKRGVIGWYDA 187
DB 101 DHVAAYVQSVYGVYQKAPAL--HSQGYTGSNVKVAIVDSIGDSHPDL--KVAGASAM 157
QY 188 VNGRSTPY--DDQGHGTHVAGTVAGTGSVNSQYIGVAPKLVGVKVLGADSGSVSTIIA 246
DB 158 VSEETNPFDQNNSHGHVAGVAAALNN-SIGVLGVAPSAISLAVKVLGADSGSGQYSWIIN 216
QY 247 GVDVYQNKXGIRVNLSTGSSQSDGTDLSQAANNAMDAGIVYCVAAAGNSGP--NT 304
DB 217 GIEMLAN--NMVYINNLSG--GPGSALKAADVAVAGVVAAGNBTGSS 269
QY 305 YTVGSPAAASKVITYGAVDSNDNIASFSSRGPETADGRLKEPVVAGVDIIAPRAGTSMG 364
DB 270 STVGYRGKPSVIANGADSSNQGRASFSSVGEELD-----VMAQVSI-----QS 314
QY 365 TPINDYTKASGTSMATPHVSGVGLIIQAPHSWTPDKYKALLETADIAPKEIAD-IA 423
DB 315 TLPGNKYGVNGTSMASPHVAGAAALILSKHNMVTQVRSLENTT-----TKLGDSPFY 369
QY 424 YGAGRNVYKA 434
DB 370 YGKGLINVOAA 380

RESULT 6
T28159
pyrolysin (EC 3.4.-.-) - Pyrococcus furiosus
C:Species: Pyrococcus furiosus
C:Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text change 09-Jul-2004
C:Accession: T28159
R:Vochohorst, W.G.B.; Eggen, R.I.L.; Geerling, A.C.M.; Plateeuw, C.; Sieren, R.J.; Vos, J.
J. Biol. Chem. 271, 20426-20431, 1996
A:Title: Isolation and characterization of the hyperthermostable serine protease, pyrolyse
A:Reference number: Z20481; MUID:96355370; PMID:8702780
A:Accession: T28159
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1398 <VOO>
A:Cross-references: UNIPROT:P72186; UNIPARC:UPI000016FD59; EMBL:U55835; NID:g1556462; PII
A:Experimental source: DSM3638
C:Genetics:
A:Gene: pls
C:Keywords: hydrolase; serine proteinase


```

QY 303 -NTYTVGSPAAASKVITVGAVDSNDNIASFSSRGPTADGRLEKPEYVAVDIIAPRASGT 361
DB 160 GNTNTIGVPAKDVSIYVAGVADNSNRASFSSVG-----AELEVMAG-----AGV 205
QY 362 SMGTPINDYITKASGTSNATPHVSGVALIIQAPSWTPDKYKTALLETADIAPKEIAD 421
DB 206 YSTYETNTYAT-LNGTNSASEHVAAGAAALILSKHPNLSASQVNRRLSTATYLG-----SS 260
QY 422 IAYGAGRVYVYKA 434
DB 261 FYYGKGLINVEAA 273

```

RESULT 10

```

Jc4802
alkaline protease (EC 3.4.21.-) precursor - Thermactinomyces sp. (strain E79)
C/Species: Thermactinomyces sp.
A/Variety: strain E79
C/Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 05-Oct-2004
C/Accession: Jc4802
R/Lee, J.K.; Kim, Y.O.; Kim, H.K.; Park, Y.S.; Oh, T.K.
Biochem. Biotechnol. Biochem. 60, 840-846, 1996
A/Title: Purification and characterization of a thermostable alkaline protease from Ther
A/Reference number: Jc4802; MUID:96261070; PMID:8704314
A/Accession: Jc4802
A/Molecule type: DNA
A/Residues: 1-384 <LES>
A/Cross-references: UNIPROT:Q56365; UNIPARC:UPI00000B2BBF; GB:U1759; NID:G1389689; PIND:
A/Experimental source: strain E79
C/Comment: This protein is thermostable.
C/Function:
A/Description: extracellular alkaline serine proteinase [validated, MUID:96261070]
C/Superfamily: Subtilisin; subtilisin homology
C/Keywords: hydrolase; serine proteinase
F/1-25/Domain: signal sequence #status predicted <SIG>
F/26-106/Domain: propeptide #status predicted <PRO>
F/107-383/Product: alkaline proteinase #status experimental <MAT>
F/134-344/Domain: subtilisin homology <SBT>
F/133-176/330/Active site: Asp, His, Ser #status predicted

```

```

Query Match 15.0%; Score 515.5; DB 2; Length 384;
Best Local Similarity 32.9%; Pred. No. 4.6e-21;
Matches 147; Conservative 77; Mismatches 150; Indels 73; Gaps 18;

QY 1 MKRLGAVTALVIVGLIA--GTALAPVKFVNRNNAVOQKNGILTPGLFKVQYQANMNQ 58
DB 1 MKRFLSVATALLLVLLAVPGTWFAA--SPASTDVPVPELIVKFDGISAQ----- 50
QY 59 EVDTVIMFGSGYGRDRAVKRLRLMGAQVYKYLIPAVAVKTKARDLLIAGMIDTGYFG 118
DB 51 --STOSTHAGYQ--AKSTIEKSKYLGFV-----VKFDGS-----VEKKIEK-YKN 90
QY 119 NTFVSGIKFIQEDYKQV-----DQATG-----VSOIGADVTNNSLIGVDSGVVAIVDTGI 170
DB 91 NPNV---EYEPRHVYHIMMTRPDLTSROMGPKQVAPQAMD-VTRSSSTYIAIVDTGV 146
QY 171 DANHPDLKGVIGYVAVNAGRSTPYDDQGHGTHVAGTGVNSQYIGVAPGAKLVGV 230
DB 147 QTHHPDLQGIYQGYVFNDSNPQDNGHGHGTHGAGIAAATVNGTIGAGMAPNMSIMPV 206
QY 231 KVLGAGSGSVSTIINGVDMVYVONKQKIGIRVYNISLSSGSSSDGTDSLISQAVNNAMDAG 290
DB 207 RVLNAGSGTMAVANGIYVAAQN---GADVISLSLG---GTSGSALQSAVQQAAMNSG 259
QY 291 IYVCAVNAAGSPNTYTVGSPAAASKVITVGAVDSNDNIASFSSRGPTADGRLEKPEYVARG 350
DB 260 AYVVAAGNS--SSSTPNYPAYYSQALNVASTDNSLSYFSNYGKWD-----VAAPG 311
QY 351 VQIIAPRASGTSMGTPINDYITKASGTSNATPHVSGVALIIQAPSWTPDKYKTALLET 410
DB 312 SNIYS-----TYLNSYASLSGTSMATPHVAGIALL--ASQGRNSQIRRAIENT 360

```

```

QY 411 ADVAPKEIDIVYAGRVNRYAIKY 437
DB 361 ADKISG---TGTYFOHGRINAYKAVNY 384

```

RESULT 11

```

JH0778
subtilisin (EC 3.4.21.62) NAT precursor - Bacillus subtilis (strain natto NC2-1)
N/Alternate names: natto proteinase; nattokinase; subtilisin BSP
C/Species: Bacillus subtilis
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Oct-2004
C/Accession: JH0778; J50601; J50517; Jc2036
R/Nakamura, T.; Yamagata, Y.; Ichishima, E.
Biochem. Biotechnol. Biochem. 56, 1863-1871, 1992
A/Title: Nucleotide sequence of the subtilisin NAT gene, aprN, of Bacillus subtilis (natt
A/Reference number: JH0778; MUID:93113095; PMID:1369081
A/Accession: JH0778
A/Molecule type: DNA
A/Residues: 1-381 <NAK>
A/Cross-references: UNIPROT:P35835; UNIPARC:UPI000000A88; GB:D25319; NID:G435439; PIND:
R/Sumi, H.
Kagaku To Seibutsu 29, 119-123, 1991
A/Title: Natto kinase and fibrinolysins.
A/Reference number: J50601
A/Accession: J50601
A/Molecule type: protein
A/Residues: 107-381 <SUM>
A/Cross-references: UNIPARC:UPI00001565DB
R/Sumi, H.; Nakajima, N.
Nippon Nogei Kagaku Kaishi 65, 1125-1127, 1991
A/Title: Studies on fibrinolytic enzymes in fermentation food.
A/Reference number: J50517
A/Accession: J50517
A/Molecule type: protein
A/Residues: 107-381 <SU2>
A/Cross-references: UNIPARC:UPI00001565DB
R/Fujita, M.; Nomura, K.; Hong, K.; Ito, Y.; Asada, A.; Nishimuro, S.
Biochem. Biophys. Res. Commun. 197, 1340-1347, 1993
A/Title: Purification and characterization of a strong fibrinolytic enzyme (nattokinase)
A/Reference number: Jc2036; MUID:94107337; PMID:8280151
A/Accession: Jc2036
A/Molecule type: protein
A/Residues: 107-381 <FUJ>
A/Cross-references: UNIPARC:UPI00001565DB
C/GeneticS:
A/Gene: aprN
A/Start codon: GTG
C/Superfamily: Subtilisin; subtilisin homology
C/Keywords: hydrolase; serine proteinase; zymogen
F/1-25/Domain: signal sequence #status predicted <SIG>
F/30-106/Domain: activation peptide #status predicted <PRO>
F/107-381/Product: subtilisin NAT #status experimental <MAT>
F/129-341/Domain: subtilisin homology <SBT>
F/138-139,170,327/Active site: Asp, Ser, His, Ser #status predicted

```

```

Query Match 14.9%; Score 513.5; DB 2; Length 381;
Best Local Similarity 38.2%; Pred. No. 5.9e-21;
Matches 139; Conservative 52; Mismatches 124; Indels 49; Gaps 13;

```

```

QY 78 VLRMGAQVYKSYKILIPAVAVKIKARDLLIAGMIDTGYRGNTRVSGIKFIQEDYKQVD 137
DB 58 VISERKGVOKQFQFYVAAATLDEKAVKEV-----KKDPVAVAYVEDH-IAHE 105
QY 138 DATSV---SOIGADVTNNSLGYDSGVVAIVDTGIIDANHPDLKGVIGMYDYNRST 193
DB 106 YAGSVPIGSIQIKPAL-HSGGTGYSNVKVAVIDSGIDSSHPLD--NVRGASFPSETN 162
QY 194 PYDD-QGHGTHVAGTGVNSQYIGVAPGAKLVGVKYLGAAGSGSVSTIAGVDMV 252
DB 163 FYQSGSHGTHVAGTIALN--SIGVLGVAPSAISLVAVKTLDSGSGQYSMIINGIEMAI 221
QY 253 QNKQKIGIRVYNISLSSGSSDGTDSLISQAVNNAMDAGIYVCAAGNSGP--NTYTVGSP 310

```


Db 222 SN----NMDVNNSLG---GPTGSTALKTIVDKAVSSGIVVAAAAGNBSGSGSTVGP 274
 QY 311 AAASKVITVGAVDSDNINIAFSRSGPTADGRLEKPEVAPGVDDIAPRAGSTSMGTINDY 370
 Db 275 AKPSTIIVAGVAVSSNORASFSSVSGELD-----VMAPEVSIQSTLPQGT----- 319
 QY 371 YTKASGTMATPHVSGVGLIILOAHPSPDKYKTLIETADIVAPKEIADIAYGAGRNV 430
 Db 320 YGAVNGTSNATPHVAGAAALLISKHPTWTMAQVRDLREISTATYLG-----NSFYGGGLIN 375
 QY 431 YVKA 434
 Db 376 VQAA 379

RESULT 12
 SUBSI
 N:Alternates: Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 04-Dec-1986 #sequence, revision 04-Dec-1986 #next change 05-Oct-2004
 C:Accession: A00972; A26116; I39778; I39779; S68012; H65586
 R:Stahl, M.L.; Ferrar, E.
 J. Bacteriol. 158, 411-418, 1984
 A>Title: Replacement of the Bacillus subtilis subtilisin structural gene with an in vitro
 A:Reference number: A00972; PMID:84212198; PMID:6427178
 A:Accession: A00972
 A:Molecule type: DNA
 A:Residues: 1-381 <STA>
 A:Cross-references: UNIPROT:P04189; UNIPARC:UPI0000142525; GB:K01988; NID:g143519; PIDN:
 R:Wong, S.L.; Price, C.W.; Goldfarb, D.S.; Doi, R.H.
 Proc. Natl. Acad. Sci. U.S.A. 81, 1184-1188, 1984
 A>Title: The subtilisin E gene of Bacillus subtilis is transcribed from a sigma37 promoter
 A:Reference number: A26116; PMID:84144862; PMID:6322190
 A:Accession: A26116
 A:Molecule type: DNA
 A:Residues: 1-155 <WON>
 A:Cross-references: UNIPARC:UPI000016E997; GB:K01443; NID:g143665; PIDN:AAA22814.1; PID:
 R:Ikemura, H.; Takagi, H.; Inouye, M.
 J. Biol. Chem. 262, 7859-7864, 1987
 A>Title: Requirement of pro-sequence for the production of active subtilisin E in Baccher
 A:Reference number: I39969; PMID:87222417; PMID:3108260
 A:Accession: I39970
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Residues: 1-156 <RKE>
 A:Molecule type: DNA
 A:Cross-references: UNIPARC:UPI000016E956; GB:M16639; NID:g143521; PIDN:AAA22744.1; PID:
 R:Renner, D.J.; Ferrar, E.; Perego, M.; Hoch, J.A.
 J. Bacteriol. 170, 296-300, 1988
 A>Title: Location of the targets of the hpr-97, sacU32(Hy), and sacQ36(Hy) mutations in
 A:Reference number: I39778; PMID:88086885; PMID:2447063
 A:Accession: I39778
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 18 <HEN>
 A:Cross-references: UNIPARC:UPI000016E7CE; GB:M19125; NID:g142527; PIDN:AAA22245.1; PID:
 R:Park, S.
 J. Bacteriol. 171, 2657-2665, 1989
 A>Title: Bacillus subtilis subtilisin gene (aprE) is expressed from a sigma-A (sigma-43)
 A:Reference number: I39779; PMID:89213955; PMID:2496113
 A:Accession: I39779
 A:Molecule type: DNA
 A:Residues: 1-13 <PAR>
 A:Cross-references: UNIPARC:UPI000016E7CE; GB:M31060; NID:g142529; PIDN:AAA22246.1; PID:
 A:Experimental source: strain W168, substrain PY79
 R:Kamal, M.; Hoegee, J.O.; Kaiser, R.; Shafgat, J.; Razzaki, T.; Zaidi, Z.H.; Joernvall,
 P. J. Biol. Chem. 263, 363-366, 1988
 A>Title: Isolation, characterization and structure of subtilisin from a thermostable Bac
 A:Reference number: S68012; PMID:96069945; PMID:7589571
 A:Accession: S68012
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 113-323 <KAM>
 A:Cross-references: UNIPARC:UPI0000172C13
 R:Kunze, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertec
 C.; Bron, S.; Brouillet, S.; Bruchli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi
 A.; Enllich, S.D.; Emerson, P.T.; Enllich, K.D.; Errington, J.; Fabret, C.; Ferrar, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Frit, C.; Fujita, M.; Fujita, Y.; Fume, S.; Gallizzi, A.; Gall
 A.; Harwood, C.R.; Henaut, H.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;
 Koester, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A.; Lander, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maasda, S.; Mausel,
 Y.; Ogawa, K.; Ogawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
 Rieger, M.; Rivolta, C.; Rocha, B.; Roche, E.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
 A.; Schleich, S.; Schroeder, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Serr,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yano, K.; Yano, K.; Yata, K.; Yoshida,
 A.; Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Zumbstein, E.; Yoshikawa, H.; Zumbstein,
 A.; Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; PMID:98044033; PMID:9384377
 A:Accession: H69586
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-26, 'V', 28-381 <KUN>
 A:Cross-references: UNIPARC:UPI000006019A; GB:Z29109; GB:AL009126; NID:g2633260; PIDN:CAF
 A:Experimental source: strain 168
 C:Comment: Secretion of subtilisin is associated with the onset of sporulation, and many
 not necessary for normal sporulation.
 C:Genetics:
 A:Gene: aprE
 A:Map position: 690-771
 A:Start codon: GTG
 C:Function:
 A:Description: catalyzes the hydrolysis of peptide bonds
 A:Note: this enzyme has broad specificity and will hydrolyze peptide amides; it prefers C
 C:Superfamily: Subtilisin, subtilisin homology
 C:Keywords: extracellular protein; hydrolase; protein digestion; serine proteinase
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-106/Domain: activation peptide #status predicted <APT>
 F:107-381/Product: subtilisin E #status predicted <MPT>
 F:129-341/Domain: subtilisin homology <SBT>
 F:138, 170, 327/Active site: Asp, His, Ser #status predicted

Query Match 14.9%; Score 512.5; DB 1; Length 381;
 Best local similarity 37.9%; Pred. No. 6, 7e-21;
 Matches 138; Conservative 53; Mismatches 124; Indels 49; Gaps 13;

QY 78 VRLMGAQVYKSYKIIPAVAVKIKARDLLIAGMDITGFGNTRVSGIKFIQEDYQVVD 137
 Db 58 VISEKGVKQKQKFNVAALTLDEKAVKEL-----KQDPSVAVEEDH-IAHE 105
 QY 138 DATSV-----SQCADTVNMSLGYDGSVVVAVDPTGIDANHPDKKRVIGWYAVNGRST 193
 Db 106 YQASVYGIQIKAPYL-HSQGYTGSNVAVVAVDSDSHPL--NVKGAASFVSEETN 162
 QY 194 PYDD-QGHGTHVAGIVAGTGSVNSQYIGVAPKALVGVKALGADSGSVSTIAGVDVV 252
 Db 163 PYDDSGSHGTHVAGTIAALAN-SIGVLGVSFBSALVAVVLDSTGSGQYSWIIINGEMAI 221
 QY 253 QNKDKYRIVNLSLSSQSSSDGTDLSQAVNNAMAGIVCVAAAGNSP--NTYTVGSP 310
 Db 222 SN----NMDVNNSLG---GPTGSTALKTIVDKAVSSGIVVAAAAGNBSGSGSTVGP 274
 QY 311 AAASKVITVGAVDSDNINIAFSRSGPTADGRLEKPEVAPGVDDIAPRAGSTSMGTINDY 370
 Db 275 AKPSTIIVAGVAVSSNORASFSSVSGELD-----VMAPEVSIQSTLPQGT----- 319
 QY 371 YTKASGTMATPHVSGVGLIILOAHPSPDKYKTLIETADIVAPKEIADIAYGAGRNV 430
 Db 320 YGAVNGTSNATPHVAGAAALLISKHPTWTMAQVRDLREISTATYLG-----NSFYGGGLIN 375
 QY 431 YVKA 434
 Db 376 VQAA 379

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2007 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2007, 22:20:17 ; Search time 192 Seconds
(without alignments)

3174.923 Million cell updates/sec

Title: US-10-800-684-1
Perfect score: 3437
Sequence: 1 MKRLGAVLALVGLAGT.....YASTYGMADYQLKAVVYG 659

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_7.2.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------|
| 1 | 2664 | 77.5 | 663 | 2 | 05J125 PYRKO |
| 2 | 2501.5 | 72.8 | 654 | 2 | 08UOC9 PYRPU |
| 3 | 982.5 | 28.6 | 561 | 2 | 08RBJ2 THETN |
| 4 | 937.5 | 27.3 | 424 | 2 | 03C855 THERT |
| 5 | 711 | 20.7 | 795 | 2 | 05NM24 9ARCH |
| 6 | 700.5 | 20.4 | 1294 | 2 | 050HM7 STRSH |
| 7 | 699 | 20.3 | 1245 | 2 | 05RL54 STRCO |
| 8 | 695.5 | 20.2 | 430 | 2 | 08ENV1 OCEIH |
| 9 | 674.5 | 19.6 | 1105 | 2 | 08KX66 STRVD |
| 10 | 669.5 | 19.5 | 442 | 2 | 031788 BACSU |
| 11 | 668.5 | 19.5 | 1139 | 2 | 082139 STRAM |
| 12 | 662 | 19.3 | 412 | 2 | 03CCT3 THERT |
| 13 | 657.5 | 19.1 | 1208 | 2 | 082B14 STRAM |
| 14 | 651 | 18.9 | 644 | 2 | 046C21 METBA |
| 15 | 649.5 | 18.9 | 442 | 2 | 05J135 GEXOK |
| 16 | 649 | 18.5 | 444 | 2 | 0651P4 BACID |
| 17 | 637 | 18.5 | 444 | 2 | 09KJ17 BACHD |
| 18 | 636 | 18.5 | 412 | 2 | 08RC68 THETN |
| 19 | 636 | 18.5 | 1102 | 2 | 095684 STRAO |
| 20 | 633 | 18.4 | 412 | 2 | 09AER6 THEHO |
| 21 | 632.5 | 18.4 | 1237 | 2 | 08GGT4 STRAZ |
| 22 | 626 | 18.2 | 435 | 2 | 08EMJ3 OCEIH |
| 23 | 623 | 18.0 | 1220 | 2 | 09LOAO STRCO |
| 24 | 619.5 | 18.0 | 1253 | 2 | 09FC06 STRCO |
| 25 | 616 | 17.9 | 369 | 2 | 03CS23 GCIOT |
| 26 | 590 | 17.2 | 1239 | 2 | 09FBZ4 STRCO |
| 27 | 564.5 | 16.4 | 379 | 2 | 09PFP4 BACLI |
| 28 | 564.5 | 16.4 | 379 | 2 | 0651P7 BACLI |
| 29 | 563.5 | 16.4 | 374 | 2 | 09P942 BACLI |
| 30 | 563.5 | 16.4 | 379 | 2 | 06BCN9 BACLI |
| 31 | 563.5 | 16.4 | 379 | 2 | 06PNN5 BACLI |

| | | | | | | |
|----|-------|------|------|---|--------------|--------------------|
| 32 | 562.5 | 16.4 | 374 | 2 | 09P943 BACLI | 09P943 bacillus 11 |
| 33 | 562.5 | 16.4 | 379 | 2 | 04PKR6 BACLI | 04PKR6 bacillus 11 |
| 34 | 559.5 | 16.3 | 379 | 2 | 053521 BACLI | 053521 bacillus 11 |
| 35 | 558.5 | 16.2 | 374 | 2 | 09P941 BACLI | 09P941 bacillus 11 |
| 36 | 557.5 | 16.2 | 379 | 1 | SUBT BACLI | P00780 bacillus 11 |
| 37 | 556.5 | 16.2 | 379 | 2 | 045259 BACLI | 045259 bacillus 11 |
| 38 | 553 | 16.1 | 453 | 2 | 03X3G0 9ACTN | 03X3G0 rubrobacter |
| 39 | 550.5 | 16.0 | 379 | 2 | 045300 BACLI | 045300 bacillus 11 |
| 40 | 549.5 | 16.0 | 1407 | 2 | 072YJ0 BACLI | 072YJ0 bacillus ce |
| 41 | 549 | 16.0 | 379 | 2 | 045301 BACLI | 045301 bacillus 11 |
| 42 | 549 | 16.0 | 1358 | 2 | 08ETM4 OCEIH | 08ETM4 oceanobacil |
| 43 | 547.5 | 15.9 | 514 | 2 | 04H6A0 9DBIO | 04H6A0 deinococcus |
| 44 | 547 | 15.9 | 382 | 2 | 045522 9BACI | 045522 bacillus sp |
| 45 | 544.5 | 15.8 | 376 | 2 | 076KL9 9BACI | 076KL9 bacillus sp |

ALIGNMENTS

| | | |
|-----------------------|---|---------------------------|
| RESULT 1 | 05J125 PYRKO | PRELIMINARY; PRT; 663 AA. |
| ID | 05J125 PYRKO | |
| AC | 05J125 | |
| DT | 15-FEB-2005, integrated into UniProtKB/TrEMBL. | |
| DT | 15-FEB-2005, sequence version 1. | |
| DT | 07-FEB-2006, entry version 7. | |
| DE | Subtilisin-like serine protease. | |
| GN | OrdereddictusNames=TK1689; | |
| OS | Pyrococcus kodakarensis (Thermococcus kodakarensis). | |
| OC | Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae; | |
| OC | Thermococcus. | |
| OX | NCBI_TaxID=69014; | |
| RP | [1] | |
| RC | NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | |
| RX | STRAIN=KOD1; | |
| RA | Published=15710748, DOI=10.1101/gr.3003105; | |
| RT | Fukui T., Atomi H., Kanai T., Matsumi R., Fujiwara S., Imanaka T.; | |
| RT | "Complete genome sequence of the hyperthermophilic archaeon Thermococcus kodakarensis KOD1 and comparison with Pyrococcus | |
| RT | genomes."; | |
| RL | Genome Res. 15:352-363(2005). | |
| CC | Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms | |
| CC | Distributed under the Creative Commons Attribution-NonCommercial | |
| CC | EMBL; AP006878; BAB5879.1; -, Genomic DNA. | |
| DR | GO; GO:0008233; F:peptidase activity; IEA. | |
| DR | GO; GO:0004289; F:subtilisin activity; IEA. | |
| DR | GO; GO:0006508; P:proteolysis; IEA. | |
| DR | InterPro; IPR007280; Pept arc bac_C. | |
| DR | InterPro; IPR00209; Pept_S8_S53. | |
| DR | Pfam; PF00082; Peptidase_S8; 1. | |
| DR | Pfam; PF04151; PFC; 2. | |
| DR | PRINTS; PRO0723; SUBTILISIN. | |
| DR | PROSITE; PS00136; SUBTILASE ASP; UNKNOWN 1. | |
| DR | PROSITE; PS00137; SUBTILASE HIS; UNKNOWN 1. | |
| DR | PROSITE; PS00138; SUBTILASE_SER; 1. | |
| KW | Complete proteome; Protease. | |
| SQ | SEQUENCE 663 AA; 70955 MW; 2CB68ACD388B90E CRC64; | |
| Query Match | 77.5%; Score 2664; DB 2; Length 663; | |
| Best Local Similarity | 75.9%; Pred. No. 5.1e-136; | |
| Matches | 505; Conservative 64; Mismatches 88; Indels 8; Gaps 4; | |
| QY | 1 MKRLGAVLALVGLAGTALAAVKKVYRNNAVQKKYGLTGTGPKVQRMMNQEV 60 | |
| DB | 1 MKKFAVVLALFLVGLMGSVLAAPQKPAVR-NVQKQKRYGLTGTGLPKVQRMSWDEV 59 | |
| QY | 61 DIVINEGSGDRDRAVKVLRMGAVQKYSYKIIIPAVVYKIKRBDLLTAGMIDTGYFGNT 120 | |
| DB | 60 STIINFNDQADKEKAVVILDFGAKIKRNYHIIPALAKIKKDKDIIITAGLMIDTGYFGNA 119 | |
| QY | 121 RVSGIKFIQEDYKQVDDAT-----SVSQIGADTVNSLGYDGSVVVAIVDTGIDANHP 175 | |

```

Db      120 QLSGVQFIOEDYVYKVAVETSLDSSAQWATMMNM-LGIDSGSITIGITIDIGIDASHP 178
Qy      176 DLKGVIVWYAVNKRSTPYDDQGHGTHVAGIVAGTGSV-NSQYIGVAPGAKLVGVYLG 234
Db      179 DLQGVIVWYVFNKKTIPYDNGHGTIVASIAAGTGAASNGKYGAPGAKLVGIVKLVN 238
Qy      235 ADGSGSVSTIIAGVWVQNKDKYGIKIVINLSLSSGSSSDGTDLSQAANNAMADGIYVC 294
Db      239 GQSGSISDIIINGVDMANQNDKGIKIVINLSLSSGSSSDGTDLSQAANNAMADGIYVC 298
Qy      295 VAAAGNSGNTYTVGSPAASKVITVGAVDSDNMTASFSSRGPTADGRLKEPVAPGVYDI 354
Db      299 VAAAGNSGNTYTVGSPAASKVITVGAVDSDNMTASFSSRGPTADNRLKEPVAPGVYDI 358
Qy      355 APRAAGTSMGTPIINDYTKASGTSMATPHVSGVALLIOAHPSWTPKVTALLETADIV 414
Db      359 AARASGTSMGPIINDYTKAAGTSMATPHVAGIALLLIOAHPSWTPKVTALLETADIV 418
Qy      415 APKEIADIAYAGRVNRYKAIKYDDYAKLTFTGSVADKGSATHTFDVSGATFTVATLYMD 474
Db      419 KPDEIADIAYAGRVNRYKAIKYDDYAKLTFTGSVADKGSATHTFDVSGATFTVATLYMD 478
Qy      475 TGSSSDIDLYLDPNNGNEVDISYTAAYGFEKVGYYNPTAGTWTVKVSYGAAANYQVDV 534
Db      479 NSGSDIDLYLDPNNGNEVDISYTAAYGFEKVGYYNPTAGTWTVKVSYGAAANYQVDV 538
Qy      535 DGSLSGSGGNPNPNPNPTPTDTQTFGTVNDVMDTSDTFMNNVNSGATKITGDLT 594
Db      539 DGSLSGSGGNPNPNPNPTPTDTQTFGTVNDVMDTSDTFMNNVNSGATKITGDLT 598
Qy      595 DTSYNDIDLYLDPNNGNEVDISYTAAYGFEKVGYYNPTAGTWTVKVSYGAAANYQVD 654
Db      599 DTSYNDIDLYLDPNNGNEVDISYTAAYGFEKVGYYNPTAGTWTVKVSYGAAANYQVD 658
Qy      655 VVYVYG 659
Db      659 KVVYVYG 663

```

RESULT 2
08UOC9_PYRFU PRELIMINARY; PRT; 654 AA.

AC 08UOC9,
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 19.
DE Alkaline serine protease.
GN Ordered locus names=PF1670;
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb P.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial license
CC EMBL; AB010265; AA181794.1; -; genomic_DNA.
DR HSSP; Q99405; IMPT.
DR BioCyc; PFUR186497; PF1670-MONOMER; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilisin-like activity; IEA.
DR GO; GO:000508; F:protease activity; IEA.
DR InterPro; IPR007280; Pept_arc_bac_C.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; Ppc; 2.

```

DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN 1.
DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR Complete proteome.
KW SEQUENCE 654 AA; 70231 MW; 1CB145A5F505DB34 CRC64;
Query Match 72.8%; Score 2501.5; DB 2; Length 654;
Best Local Similarity 72.8%; Pred. No. 3.2e-127;
Matches 484; Conservative 69; Mismatches 93; Indels 21; Gaps 8;
Qy      1 MKRLGVVIALVYVGLAGTALAAAPYKPY--VNNNAVQKNGILRGFKYQVRNNNQ 58
Db      1 MKRLKALIVILVYVGLVGVSVAAAPKRYEQVRN--VEKNGILRGFKYQKLNPNB 57
Qy      59 EVDVTIMPSGYGDRDAVKVRLMGAVKYSKIIPAVAVKIKAQDILLIAGMIDTGYG 118
Db      58 EISTVIVFENHREKEIAVVLLEMGAKVYVYVHIIPALAADLKVRDLVLSGL--TG--G 113
Qy      119 NTRVSGIKRFOEDYKYQVNDV-----TSVQIGADTVNMSLGIDSGVVAIVDTGIDAN 173
Db      114 KALDSGVKRFIOEDYKVTVAELLEGDSAQWATYVNM-LGIDSGSITIGITIDIGIDAS 172
Qy      174 HPDLKGVIVWYAVNKRSTPYDDQGHGTHVAGIVAGTGSV-NSQYIGVAPGAKLVGVYK 232
Db      173 HPDLQGVIVWYVFNKKTIPYDNGHGTIVASIAAGTGAASNGKYGAPGAKLVGIVK 232
Qy      233 LGADSGSVSTIIAGVWVQNKDKYGIKIVINLSLSSGSSSDGTDLSQAANNAMADGI 292
Db      233 LGADSGSVSTIIAGVWVQNKDKYGIKIVINLSLSSGSSSDGTDLSQAANNAMADGI 292
Qy      293 VCAAGNSGPNYTVGSPAASKVITVGAVDSDNMTASFSSRGPTADGRLKEPVAPGV 352
Db      293 VVAAGNSGPNYTVGSPAASKVITVGAVDSDNMTASFSSRGPTADGRLKEPVAPGV 352
Qy      353 IAPRASGTSMGTPINDYTKASGTSMATPHVSGVALLIOAHPSWTPKVTALLETAD 412
Db      353 IIAARASGTSMGPIINDYTKAAGTSMATPHVAGIALLLIOAHPSWTPKVTALLETAD 412
Qy      413 IYAPKEIADIAYAGRVNRYKAIKYDDYAKLTFTGSVADKGSATHTFDVSGATFTVATLY 472
Db      413 IYKPEIADIAYAGRVNRYKAIKYDDYAKLTFTGSVADKGSATHTFDVSGATFTVATLY 472
Qy      473 WDTGSSSDIDLYLDPNNGNEVDISYTAAYGFEKVGYYNPTAGTWTVKVSYGAAANYQVD 532
Db      473 WMANSDIDLYLDPNNGNEVDISYTAAYGFEKVGYYNPTAGTWTVKVSYGAAANYQVD 532
Qy      533 VSDGSLSGSGGNPNPNPNPTPTDTQTFGTVNDVMDTSDTFMNNVNSGATKITGDL 592
Db      533 VSDGSLSGSGGNPNPNPNPTPTDTQTFGTVNDVMDTSDTFMNNVNSGATKITGDL 592
Qy      593 TFDTSYNDIDLYLDPNNGNEVDISYTAAYGFEKVGYYNPTAGTWTVKVSYGAAANYQVD 652
Db      588 VDTSYNDIDLYLDPNNGNEVDISYTAAYGFEKVGYYNPTAGTWTVKVSYGAAANYQVD 647
Qy      653 KAVVYVYG 659
Db      648 TAKVYVYG 654

```

RESULT 3
08RBJ2_THETN PRELIMINARY; PRT; 561 AA.

AC 08RBJ2,
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 21.
DE Subtilisin-like serine protease.
GN Name=AprE2; Ordered locus names=TR0824;
OS Thermobacterium tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermobacteriales;
OC Thermobacteriaceae; Thermobacter.
OX NCBI_TaxID=119072;


```

OY 360 ATPHVSQVAILLOAHPSMTDPDKVTALIEADIVAPKEIADIVAGAGRVVYKAIKDD 439
DB 246 ATPFVAGTVALMLSANINILAPLAKNIIMTTAKSGPPS-KNIDYGVGRDLAYEALIKTAG 304
OY 440 -----YAKLFTGSGVADKSGATITFDVSGATFTATLY----WDGSSDIDL 482
DB 305 NFTGNTISVPHNYAKESLPSS---RYSIDIWTFVTDTISYPIATFTIIPDMANNPFDI 361
OY 483 YLYDPNGNEVDVSYTAYAYGFEKGYVNPATGTMATVYSYKGAANYQVDV-YSDGSLSSQ 541
DB 362 YLYDPTGLVK-SSTGTGTGTITITITITITITITITITITITITITITITITITITIT 420

RESULT 5
O5NM24_9ARCH PRELIMINARY; PRT; 795 AA.
ID O5NM24_9ARCH
AC O5NM24_9ARCH
DT 04-JUN-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 1.
DE Alkaline serine protease.
GN ORFNames=orf17;
OS uncultured archaeon.
OC Archaea; environmental samples.
OX NCBI_TaxID=115547;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=16329940; DOI=10.1016/j.femsec.2004.12.004;
RA Etkel C., Kemnitz D., Kube M., Rieke P., Chin K.-J., Dedysh S.,
RA Reinhardt R., Conrad R., Liesack W.;
RT "Retrieval of first genome data for rice cluster I methanogens by a
RT combination of cultivation and molecular techniques.";
RL FEKS Microbiol. Ecol. 53:187-204(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
EMBL: CR626856, CA04782.1, -, Genomic DNA.
GO: GO:0008233; F:peptidase activity; IEA.
GO: GO:0042802; F:protein self binding; IEA.
GO: GO:0004289; F:subtilase activity; IEA.
GO: GO:0043086; P:negative regulation of enzyme activity; IEA.
DR InterPro: IPR011964; Beta_rpt_yetn.
DR InterPro: IPR000209; Pept_S8_S53.
DR InterPro: IPR010259; Prot_inh_S8A.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN; 1.
DR TIGRPFAM: TIGR02276; beta_rpt_yetn; 3.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; UNKNOWN_1.
KW Protease.
SQ SEQUENCE 795 AA; 82723 MW; 629536B8B3D16B88 CRC64;

Query Match 20.7%; Score 711; DB 2; Length 795;
Best Local Similarity 31.6%; Pred. No. 4.1e-30;
Matches 194; Conservative 92; Mismatches 207; Indels 120; Gaps 16;

OY 57 NOEVDVIVFGSGDRDAVKVLRMGAVKYSYKIIIPVAVKIKARDDLILAGMIDICY 116
DB 31 NKGQASLISQGAASGKATRLRVKAGKVKNNIIDAVALDPDSEVAALAKAPD--- 87
OY 117 FGNTRVSGIKFTQEDYKVVADATSVSQTADTVNSL-GYDGSVVVAIVDTGIDANHP 175
DB 88 -----VESVVRDSIAVADLDEVVYHNAVASVASTVGVGTGKGVNVSVIDSGIDAHHP 139
OY 176 DLKGVITGVYDVAVNGSTPYDQGHGHTVAGIVAGTGSVNSQYI-GVAPGAALVGVKVLG 234
DB 140 DLAKGVILMKDPLNDATPYDDFGHGTFAAGIIASGANSGEVGLVADASTLFEVKTILN 199
OY 235 ADGSGSVSTIIAGVDVWVQNKDKYGIARVILNLSGSSQSSDGTDSLQAVNNAMWDGIYVC 294

```

```

DB 200 PSGTAYVSDIILAAIDSVQNH-----ADVLSMSLSNPTH---IQALDAVNAANEVWVAV 252
OY 295 VAAAGSGPNTYTVGSPAAASKVITTVAGVDSNDNIASFSSRGPTADGRILKPEVAPGVDI 354
DB 253 CSAAGNTGPKYKGSIRCRGDSPPDVIAVGSVMSBRLSSFSRGPTRIDRIEDIVAVEYVI 312
OY 355 APPASGTSMTPIINDYTTAASGTSNATPHVSGVAILLOAHPSMTDPDKVTALIEADIV 414
DB 313 SYRSSGTSWGNPILGQYCYASGTSAACPQVSAASAILLOANGSLTFPEEKIDVILRNT--- 369
OY 415 APRKADI-----AYGAGRVVYKAIKDDYAKLFTGSGVADKSGATITFDVSGATFTAT 470
DB 370 ---YHLSDTYPENQSGKRLINIAL--NEVLOVTPPTPTPTPTPTPTPTPTPTPTPTAT 425
OY 471 LWDGSSDIDLVLDPNGNEVDVSYTAYAYGFEKGYVNPAT---GTWTVKVSYKGAAN 527
DB 426 -----FLATPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 440
OY 528 YQVDVVSQDLSQSGGAGNPENPNPTPT-----TDTQFTGSGVNDY 570
DB 441 -----TVPTPTATPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 491
OY 571 WDTSDFTTN-----VNSGATKITGD-LTFDTSYNDLVLDPNGNLVDRSTSSNS 621
DB 492 PPKSKTFVNSGNTVSVVDQSTNTVTGIIIGSRPYG-----IVNSPDGSLV--YVAVEG 545
OY 622 YEHVERANAPGT 634
DB 546 LNRVAIISPANNT 558

```

RESULT 6

O5OHM7_STRSH

```

ID O5OHM7_STRSH PRELIMINARY; PRT; 1294 AA.
AC O5OHM7;
DT 07-JUN-2005, integrated into UniProtKB/TrEMBL.
DT 07-JUN-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Probable secreted peptidase.
OS Streptomyces sphaeroides.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=195949;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NCIMB 11891;
RX PubMed=15870333; DOI=10.1128/AEM.71.5.2452-2459.2005;
RA Bustaquito A.S., Gust B., Galm U., Li S.-M., Chater K.F., Heide L.;
RT "Heterologous Expression of Novobiocin and Clorobiocin Biosynthetic
RT Gene Clusters.";
RL Appl. Environ. Microbiol. 71:2452-2459(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NCIMB 11891;
RA Steffensky M., Li S.-M., Heide L.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
EMBL: AY227005; AAP48601.1, -, Genomic DNA.
GO: GO:0008233; F:peptidase activity; IEA.
GO: GO:0004289; F:subtilase activity; IEA.
GO: GO:0006508; P:proteolysis; IEA.
DR InterPro: IPR001337; PA.
DR InterPro: IPR000209; Pept_S8_S53.
DR Pfam: PF02225; PA; 1.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
SQ SEQUENCE 1294 AA; 135576 MW; 804C7F9A0DDEB896 CRC64;

```


Query Match 20.4%; Score 700.5; DB 2; Length 1294;
 Best Local Similarity 45.6%; Pred. No. 2,8e-29;
 Matches 160; Conservative 53; Mismatches 117; Indels 21; Gaps 9;

93 IPAVAVKIKARDL-----LLIAGMIDTGYFGNTRVSGIKFIOEDYKQVDDATVSQIG 146
 165 IDARAVRVPADLGFAMKQAPAG-----PSINSAALDAVPKWLDRVASIDRSTPAQIG 220
 147 ADTWNSLIGYDGGVVAIVDTGIDANHPDLKGIWTDVANGKSTPTDDGCHGVAVG 206
 221 ABDVWKS-GLRGERVAVVATLDGADQTHDPLAGRIAAAD-PSGSGTADGSGHGVAS 278
 207 IYAGTGSVNV-SQYIGVAPGAKLVGVVVLGADGSGVSTIAGVDVWQKDKKXGIRVIL 265
 279 IYGGSGKASGGRQGAAPAEELMIGVLDGDDGSGSYIAGMEMAAK-----GAEVVM 334
 266 SLGSSQSDGDTSLGQAVNN-AMDAGIVVCVAAGNSGPTTYVSGPAAASKVITVGVADS 324
 335 SLGSDAPSDGTPMSLAVNELSESSGALFVVAAGNSGPGSGTIGSGAADALITVGAVD 394
 325 NNINIAFSRRGP-TADGRLEPVAVPGVDI IAPRASGTMGTPINDYTKASGTMATG 383
 395 DDLAEFSSRGPRSGDEAVKPDVTAPGVI VAAARATGTMGDPVDDGYTAAGSTMATG 454
 384 VSGVALLIQAHPSTPDVKYKTALIEFDIIVAPKEIADIAYGAGRVNYKA 434
 455 VAGAAALLAQRHPDWSAAQLKDALVSTARTIAQQVTE--QGGGRIDILAA 503

DB

RESULT 7
 Q9RL54_STRCO PRELIMINARY; PRT; 1245 AA.
 ID Q9RL54;
 AC 01-MAY-2000, integrated into UniProtKB/TrEMBL.
 DT 01-MAY-2000, sequence version 1.
 DT 07-FEB-2006, entry version 21.
 DE Probable secreted peptidase.
 OS OrderedLocustNames=SC00432; ORFNames=SCF51A.10;
 OC Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.
 NX NCBI_TaxId=1902;
 [1]
 RA NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleeer H.,
 Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 Cronin A., Fraser A., Goble A., Hidalgo J., Hornby T., Howarth S.,
 Huang C.-H., Kleeer T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
 Rabinovitch R., Rajandream M.A., Ruberford K.M., Ruter S.,
 Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
 Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)."
 RL Nature 417:141-147(2002).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC EMBL: AL939105; CAB5662.1; -; Genomic_DNA.
 DR HSSP: P00782; 1SUR.
 DR BIOCyc: SC01902; SC00432-MONOMER; -;
 DR GO: GO:0008233; F:peptidase activity; IEA.
 DR GO: GO:0006508; P:proteolysis; IEA.
 DR InterPro: IPR003137; PA.
 DR InterPro: IPR000209; Pept_S8_S53.
 DR InterPro: IPR010221; VCSB-
 DR Pfam: PF02225; PA; 1.

DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PR00723; SUBTILISIN.
 DR TIGRFAMs: TIGR01965; VCSB repeat; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KW Complete proteome.
 SQ SEQUENCE 1245 AA; 130896 MW; 74EE92DB9CA1DE60 CRC64;

Query Match 20.3%; Score 699; DB 2; Length 1245;
 Best Local Similarity 41.1%; Pred. No. 3.2e-29;
 Matches 175; Conservative 67; Mismatches 134; Indels 50; Gaps 13;

72 RDAVAVKRLMGAQV-----KYSYKIIIPAVAVKIKARDLLIAGMIDTGYGNTFR 122
 157 RDAVRELDSTIDARAVRDEAMELGRFMKRLVPAGQRAKA-----A 197
 123 SGIRKIQEDYKQVDDATVSQIGADTWNSLIGYDGGVVAIVDTGIDANHPDLKGIY 182
 198 AATPRVWLDGRVGAISIDRSTAQIGAPDVW-SAGYRGDGKVAIVLDTGADQSHDLAGRVA 256
 183 GMYDAVNGRSTPYDDQGHGVAGI VAGTSVN-SQYIGVAPGAKLVGVKVLGADGSGSV 241
 257 AAKD-PSGSSGNTDVFGEHGVASIVGSGAASGSGRQGVAPAAKLVGKVLGDDGFSB 315
 242 STIAGVWVWQNKQKRYGIRVNLISGSSQSDGDTSLGQAVNN-AMDAGIVVCVAAGNS 300
 316 SGVINGMEMAADO-----GADVVMNLGSSGATDTPMSQALNLSRRTGTLFVVAAGNE 371
 301 G-PNTYVGSPPAAASKVITVGAIVDSNDNIASFSSRGP-TADGRLEPVAVPGVDIIAPRA 358
 372 GEGQPRIVGSGPAAAPALVGVADVDDSLAPSSSRGPRGDAVXPDVAPGVGIIAARA 431
 359 SGTSMGTPINDYTKASGTMATPHVSGVGAIIILAHPEWTPDKYKTALIEFDIIVAPKE 418
 432 AGSAGMDPVDEHYTASGTMATPHVAGAAALLAQHPMTGAQLDALISTAVYDQK 491
 419 IADIVYAGRVVVAIKYDPAKLTFTGSVADKSGATHTFPVSGATFVATLTYMDGSS 478
 492 VTE--QGGGRIVRAA-----GIGAVTATGTL-----VMGPFISRDPETPSRVRY-TNS 539
 479 DIDVLT 484
 540 DEDVTL 545

DB

RESULT 8
 Q8ENV1_OCEIH PRELIMINARY; PRT; 430 AA.
 ID Q8ENV1;
 AC Q8ENV1;
 DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
 DT 01-MAR-2003, sequence version 1.
 DT 07-FEB-2006, entry version 19.
 DE Intracellular alkaline serine proteinase.
 OS OrderedLocustNames=OB2375;
 GN Oceanobacillus iheyensis.
 OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
 NX NCBI_TaxId=182710;
 [1]
 RA NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=HTB831 / DSM 14371 / JCM 11309;
 RX MEDLINE=22220767; PubMed=12235376; DOI=10.1093/nar/gkf526;
 RA Takami H., Takaki Y., Uchiyama I.;
 RT Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
 RT Ridge and its unexpected adaptive capabilities to extreme
 RT environments."
 RL Nucleic Acids Res. 30:3927-3935(2002).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC EMBL: BA000028; BA014331.1; -; Genomic_DNA.
 DR HSSP: Q99405; 1MPT.
 DR BIOCyc: OIHE182710; OB2375-MONOMER; -;

DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR InterPro; IPR000209; P:proteolysis; IEA.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILASIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Complete proteome; Hydrolase; Protease; Serine protease.
 KM SEQUENCE 430 AA; 45839 MW; 6D09A99B8C1310F CRC64;

Query Match 20.2%; Score 695.5; DB 2; Length 430;
 Best Local Similarity 47.0%; Pred. No. 1.3e-29;
 Matches 155; Conservative 50; Mismatches 104; Indels 21; Gaps 8;

QY 125 IKFTQEDVKVQVDDATVSQIGADTVNMSLGVDGSGVVAIVDTGIDANHPDLKGVIGM 164
 DB 107 IEKLYDRKVPFLSDTSSINADVLKES-GLTQSGTIAVIDGTHP-HEDEGRITGF 164
 QY 185 YDANVNGSTPYDDQGHGTHVAGVAGTGSV-NSQYIGAPQAKLVGVKVLGADSGSVST 243
 DB 165 ADFYKQGTPEYDDNGHGTGACGDAAGCALSDQYQGPAPDANLVGVKVLKTKGSSGLST 224
 QY 244 IIAQVWVQNKDKYGRVNLSTGS--SQSSQSTDSLQAVNNAMPAGIVCVAAAGNSG 301
 DB 225 VIESIDICIQOSKTNINILSLSGSDATPEAG-DEVVNAVETAMNGVVCVAAAGNSG 283
 QY 302 PNTYVSGPAAASKVITVGAVDNSN-----DNIAFSPSRPTADGLKPEVAVAPVDII 354
 DB 284 PGDKTVSGPSIPVITVGAADNNTARSDDSVAEFSRGPITDGLTKNMLTPVDIV 343
 QY 355 APRASG-----TSMGTINDYTTKASGTSMAHPVSGVALIIQAPHSWTPDKVKTALIE 409
 DB 344 SLRPSGFDIKTNKSAKRVGNYISLSGTSMAHPICAGIVAQIQQSDSLTPNQVKELME 403
 QY 410 TADIVAKELADIYVAGRVNVVKAIKYDD 439
 DB 404 ACQDILG---SPNVQAGYMANANLIVINE 430

RESULT 9
 Q8KH6 STRVD
 ID Q8KH6 STRVD PRELIMINARY; PRT; 1105 AA.
 AC Q8KH6
 DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.
 DT 01-OCT-2002, sequence version 1.
 DT 07-FEB-2006, entry version 16.
 DE 1,4-dihydropyridine enantioselective esterase precursor.
 GN Name-dhpA;
 OS Streptomyces viridosporus.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=67581;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=A-914;
 RX MEDLINE=22034940; PubMed=12039725;
 RA DOI=10.1128/AEM.68.6.2716-2725.2002;
 RA Aribea A., Matsufuji M., Nakashima T., Dobashi K., Ieshiki K.,
 RA Yoshioka T., Yamada S., Momose H., Taguchi S.;
 RA "Streptomyces serine protease (DHP-A) as a new biocatalyst capable of
 RT forming chiral intermediates of 1,4-dihydropyridine calcium
 RT antagonist.";
 RL Appl. Environ. Microbiol. 68:2716-2725(2002).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonCommercial License
 CC EMBL; AB007809; BAC00500.1; -; Genomic_DNA.
 DR HSSP; P00782; 2SRT.
 DR MEROPS; S08.069; -;
 DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis; IEA.
 DR InterPro; IPR000209; Pept_S8_S53.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILASIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Hydrolase; Protease; Serine protease; Signal.
 FT SIGNAL 1 204
 KM SEQUENCE 1105 AA; 114126 MW; 6AB34D8870579AB CRC64;

Query Match 19.6%; Score 674.5; DB 2; Length 1105;
 Best Local Similarity 30.8%; Pred. No. 5.8e-28;
 Matches 219; Conservative 85; Mismatches 273; Indels 133; Gaps 27;

QY 22 LAAPVYVNRNNAVQKATGLTPGLFKVQRMNQEVDTYIMGSGYDRDRAVKVRL 81
 DB 97 LPDAARLVAASGLDRRLFDITELG--KAATNSQKQGLKIV--GYQAAARAARAEVR 151
 QY 82 MGAOVKYSKTIIPAVAVKIKARDLLIAGMIDTVGFGNTRVSGIKFQEDYKVQVDDATS 141
 DB 152 EAGELRRITLSLNADAVRTPHED--ASELMADVTNGDRTASGIAHWLWDGVRAALDT 208
 QY 142 VSQIGADTVNMSLGVDGSGVVAIVDTGIDANHPDLKGVIGVDAVNGRSTP--YDDQ 199
 DB 209 VQIGAPKAM-SAGYDGKGVKIAVLDTVGDSHPDLKGRVTA--SKNFTAPAGADKVG 264
 QY 200 HGTVAAGVAGTGSVN-SQYIGAPQAKLVGVKVLGADSGSVSTIAGVDVQNKDY 258
 DB 265 HGTVAASIAAGTGAOSKGVKVAAPGAILNGKVLDDSGFGDSQIILAGMEWAAAQ---- 320
 QY 259 GIRVNLSTGSSQSSQSDGTDSLQAVNN-AMDAIVCVAAAGNSGPNYTVSGPAAASKVI 317
 DB 321 GADVNNMSLGMDTPE-IDPLEAAYDKLSAEKGV.PALAAAGNGB--SIGSPGADAL 377
 QY 318 TVGAVDSDNINIASFSSRGP-TADGLKPEVAVAPGVDTIAPRASGSMGTPIND--YYTK 373
 DB 378 TVGAVDSDKXKLADFSSTGPRLDGAIKPDVTAAPVDITAAAGNDIQEVEGEPAGVMT 437
 QY 374 ASGTSMAHPVSGVALIIQAPHSWTPDKVKTALLETADIVAPKELADIANAGRVNYK 433
 DB 438 ISGTSMAHPVAGAAALIKQHPMTSALKAL--TGSTKGK-YTFPEQSGSIGIADK 494
 QY 434 AIK-----YDDYAKLTFGSAVDKSGATH 457
 DB 495 ALQGVIVADPVSVSRGVQGWHTDEPVTXQLTYNLGTQVTLKLTSTATDPKGAAPA 554
 QY 458 TFDVSGATFVATLTYMTGSSDIDLYLYDPNGNEVDYTAIYGFEEKVGYNPTAGTWT 517
 DB 555 GFTLGATTVVPA--GGSASVDMTADTRLGTVDGAYSAYVATGGQTVRTAAAVQ 611
 QY 518 KVSF-----KGAANYQVDVSDSLSGSGGNPNPNPNPTPTTDTQTFTGS 566
 DB 612 EVESIDTVVRIHGRDKPTHTLTLIGVAGLSGGRVG--APATDTATRLP 662
 QY 567 VNDY-----W-----DTSDFTMNVNSGATKTGDLTF-DTSYND 600
 DB 663 KGYLVDSWIAKDRGLKGLDVLVQPLSTYTKOT-TLTDARTK-AAADITVDDPKAPK 720
 QY 601 LDV--YYIDNGNLVDRSTSSNVEHYEVANPAP-----GTWT 636
 DB 721 LSAITGYTYDTAG--IGIGASMESFADVRMAHLGPAPGIRQTWNGWT 768

RESULT 10
 O31788 BACST
 ID O31788 BACST PRELIMINARY; PRT; 442 AA.
 AC O31788;
 DT 01-JAN-1998, integrated into UniProtKB/TrEMBL.
 DT 01-JAN-1998, sequence version 1.
 DT 07-FEB-2006, entry version 31.
 DE Alkaline serine protease.

GN Name:apix; OrderedLocustNames=BSU17260;
 OC Bacillus subtilis.
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
 RA Kusnet F., Ogasawara N., Moszer I., Albertin A.M., Allion G.,
 RA Azevedo V., Besterio M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Bouriez R., Boutsier L., Brans A., Braun M., Brignelli S.C., Bron S.,
 RA Brouillet S., Brunsch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.-K., Codani J.-J., Comercon I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Duesterhoeft A., Ehrlich S.D., Emmerson P.T.,
 RA Entian K.-D., Errington J., Fabret C., Ferrari E., Fougard D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gilm S.-Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holteppel S., Hosono S., Hulio M.-F., Itaya M.,
 RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaere-Blanchard M.,
 RA Klein C., Kobayashi Y., Koester P., Koningsstein G., Krogh S.,
 RA Kuma M., Kurita K., Lapidus A., Iardinois S., Laber J.,
 RA Lazarevic V., Lees S.-M., Levine A., Liu H., Maeda S., Manuel C.,
 RA Lejeune C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
 RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
 RA Park S.-H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
 RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
 RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
 RA Rose M., Sadie Y., Sato T., Scanlan E., Schlecht S., Schroeder R.,
 RA Scoffone F., Seliguchi J., Sekowska A., Seror S.J., Serron P.,
 RA Shin B.-S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
 RA Takemaru K., Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P.,
 RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
 RA Vassarotti A., Varti A., Wambutt R., Wedler E., Wedler H.,
 RA Wetzinger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
 RA Yaumoto K., Yata K., Yoshida K., Yoshikawa H.-F., Zumbstein E.,
 RA Yoshikawa H., Darchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256 (1997).
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonCommercial license
 CC
 CC
 CC EMBL; Z99113; CAB13610.1; -; Genomic_DNA.
 DR PIR; A69587; A69587.
 DR HSP; O99405; IMPT.
 DR Biocyc; BSUB1423;BSU1727-MONOMER; -;
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis; IEA.
 DR InterPro; IPR000209; pept_88_S53.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISTIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SSR; 1.
 KM Complete proteome; Hydrolase; Protease; serine protease.
 SQ SEQUENCE 442 AA; 47906 MW; B96446ABE7BMDP2 CRC64;

```

Db      234  EGVMKCQYNEDNDDEPIDIMSMISGQALRYDHQEDPLVRAVEAMSAGIVVCVCAACN 293
Qy      300  GGPNTVTVGSPAAASKVITGAVDSN-----DNIAFPSSRGFPADGRLKEVVAAGVD 352
Db      294  SGPDSPQRTIASPGVSEKVTITGVALDDNNRTASSDDDTVAFFSSGCPVYVGKEKEDILAPGVN 353
Qy      353  IIAARASGT-----SMGTPINDYTKAGCTSMATRHVSGVALLIIQAHPSWTDPDKYKTL 407
Db      354  IISLRSPSYIDKLQKSSRSVQYFTWMSGTSWATPICAGIALILQONPDLTPEVEKELL 413
Qy      408  IETADIVAPKEIADIAYGAGGVNYYKAI 435
Db      414  KNQTD--KMKDEDPNITYGAGVNAENSV 439

RESULT 11
082139 STRAW
ID 082139_STRAW PRELIMINARY; PRT: 1139 AA.
AC 082139
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Putative subtilisin-like protease.
GN OrderedLocustNames=SAV3319;
OS Streptomyces avermiltis.
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxId=33903;
[1]
RN RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RT RT Sakaki Y., Hattori M., Omura S.;
RT RT "Complete genome sequence and comparative analysis of the industrial
microorganism Streptomyces avermiltis.",
RL RL Nat. Biotechnol. 21:526-531(2003).
[2]
RN RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.21143198;
RA RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RT RT Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT RT "Genome sequence of an industrial microorganism Streptomyces
avermiltis: deducing the ability of producing secondary
metabolites.",
RL RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
CC CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC CC Distributed under the Creative Commons Attribution-NonDerivs License
CC CC -----
CC CC EMBL, BA000030, BAC71030.1, -, Genomic_DNA.
CC CC HSSP, Q99405, IMPT.
DR DR MEROPS, S08_069, -.
DR DR BIOCYG, SAVE227882, SAV3319-MONOMER, -.
DR DR GO: GO:0008233, F:peptidase activity; IEA.
DR DR GO: GO:0004289, F:subtilase activity; IEA.
DR DR GO: GO:0006508, P:proteolysis; IEA.
DR DR InterPro, IPR002860, Glyco_hydro_BNR.
DR DR InterPro, IPR000209, Pept_S8_553.
DR DR Pfam, PF02012, BNR, 2.
DR DR Pfam, PF00082, Peptidase_S8, 1.
DR DR PRINTS, PRO0723, SUBTILISIN.
DR DR PROSITE, PS00136, SUBTILASE_AAP, 1.
DR DR PROSITE, PS00137, SUBTILASE_HIS, 1.
DR DR PROSITE, PS00138, SUBTILASE_SSR, 1.
DR DR Complete proteome; Protease.
SQ SQ SEQUENCE 1139 AA; 118193 MW; EC3B6D234FA94FA0 CRC64;

```

Matches 222; Conservative 95; Mismatches 261; Indels 109; Gaps 28;

QY 29 VYRNAVVOQKNYGLTPGLFKYKQGMNN-----OEVDYIMFGSYGRDRAVYKRLMG 83
 DB 124 VYPAAMAMIAATGKLDQRLF-DYTELNSATRKAKQKQKLYVGKYGASAKSDVRAG 182
 QY 84 AOVKSYKIPAVAVKIKARDLLIAGMIDYGFQNTVSGIKFOEDYKQVVDATSVS 143
 DB 183 T-LRRSLKSLNADAVQTPRTD---TAEIMDAVTMGDAPASGVAHVWLDGTRKASLDKSV 238
 QY 144 QIGADTVNNSLGYDSSGVVAVLDVGDIDANHPDLKSKYIGYDANV--GRSTPYDDQHG 201
 DB 229 QIGAPTM-AGGYDCKGVKIALVDGVDATPDLKDQVA---EKNKSPAAADADHFGHG 294
 QY 202 THVAGIVAGTGS-VNSQYIGVAPGAKIVGVYVLAGDSGSYSTIAGVWVYONKDKXI 260
 DB 295 THVASIAGTGAKSNGKYGVAFGATLNGKVLVDGSDGSGLLAGEMAAG---GA 350
 QY 261 RVINLSLSSSSSDGTSLQAVNN-AMDAQIVVCVAAGNSP-NTYVSSPAAASKVIT 318
 DB 351 DVNLSLGGGDTPE-IDPLEAEVVKLSEKGIILFAIAGNGEFGEGOTIGSPGSAADALT 409
 QY 319 VGADVSDNINIASFSSRGPTADGRLEKPEVAVARGVDIIAPRASGSMGPIN---DYTKAS 375
 DB 410 VGAVNDSDKLASFSSRGGLDGAIKPDTAPGVDTAAAPGSVIDEGVGQKPDGYLTIS 469
 QY 376 GTSMATPHVSGVALLIOAPSPWTPDKVKTALLETADIVAPKEIADIYAGAGRVYVYKAI 435
 DB 470 GTSMATPHVAGAAAILKQHPHNSFAELKGLTGA---KQKGYTPQSGSRIADVKA 526
 QY 436 KYDDYAKLTFTGSVADGSAAT-----HTFDVSGATFYVATL-YMTGSSDIDLVI--- 484
 DB 527 K-----QSVIANPNSVSPGICQMPHTDCKP---VTQQLRYRLGTSVTLNLAST 573
 QY 485 -YDPNGNEVDVSYTYAVYGFKEVGYNPTAGITWYKVVSYGAANYQVDVSDSLG---- 539
 DB 574 ATNPKG-----VAAPSGFPLGATKYTPVAGGASADF---TYNTKLGSTTDGAYSAVYT 625
 QY 540 QSGGAGN-----NPNPNPTPTTDTOTFTGSSVNDVY-----DTS 574
 DB 626 ATGGQGTQVTRTAAAVQREVSVDYTLKHIDRGKPAVAVSYDLTGVSGLADKFPAPYDAS 685
 QY 575 DTFPMNVNSGATKLTGDLTFD---TSYNDLILYLDVDPNGNLVDRST---SSSYEHVEY 627
 DB 686 GTYKVRPKNFILNASLFPADPEDFTKAD---WIAQPKLSVTKYTVYDAKAKAPVDI 742
 QY 628 ANPAGTWTFLVYAVSYTYGADYQAK 654
 DB 743 TVPDKG-----AKSAFASPDYTVEA 762

RESULT 12
 O3CCT3 THEET PRELIMINARY; PRT; 412 AA.
 AC O3CCT3;
 DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
 DT 22-NOV-2005, sequence version 1.
 DE 07-FEB-2006, entry version 4.
 DE Peptidase S8 and S53, subtilisin, kexin, sedolisin.
 GN ORFNames=Tech39DRAFT_1862;
 OS Thermoanaerobacter ethanolicus ATCC 33223.
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
 OC Thermoanaerobacteriaceae; Thermoanaerobacter.
 NC NCBI_TaxID=340099;
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 33223;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler J.C., Glavina T.,
 RA Hammon N., Israeli S., Plunk S., Richardson P.,
 RT "Sequencing of the draft genome and assembly of Thermoanaerobacter
 ethanolicus 3323".
 RL Submitted (SEP-2005) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 33223;
 RG US DOE Joint Genome Institute (JGI-ORNL);
 RA Laximer F., Land M.;
 RT "Annotation of the draft genome of Thermoanaerobacter ethanolicus
 3323".
 RL Submitted (OCT-2005) to the EMBL/GenBank/DBJ databases.
 CC CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonCommercial License
 CC -----
 DR EMBL; AA001000019; E064328.1; -; Genomic DNA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilisin activity; IEA.
 DR GO; GO:0006508; P:proteolysis; IEA.
 KM Hydrolyase; Protease; Serine protease.
 SQ SEQUENCE 412 AA; 44421 MW; CEBBF0782716EF8 CRC64;

Query Match 19.3%; Score 662; DB 2; Length 412;
 Best Local Similarity 37.8%; Pred. No. 8.3e-28;
 Matches 166; Conservative 72; Mismatches 155; Indels 46; Gaps 14;

QY 13 LVGLAGTALAAPVYRVNNAVVOQKNYGLTPGLFKYKQGMNNOEVDYIMFGSYGRD 72
 DB 3 LIPMFPSKIA---KEIYVSNKIDKR---LKKALYLR-----SECVFVLY-SHLSY 48
 QY 73 DRAVYKRLMGAOVKSYKIPAVAVKIKARDLLIAGMIDYGFQNTVSGIKFOEDY 132
 DB 49 DLKEKIEKGSGSIKPELPIIKAMSVNLPCLDKLKPFA-----TLKIHITADS 97
 QY 133 KQVVDATSVSQIGADTVNNSLGYDSSGVVAVLDVGDIDANHPDL---KQVYGYDANV 189
 DB 98 AVKLDQYIATQEIARNA-NDLGYTGKGVITLFTDGL-YPHPDFTPKRNRIAVHDVN 155
 QY 190 GRSTPYDDQGHGTHVAGIYAGTG-SVNSQYIGVAPGAKLVGKVLGAGSGSSVSTIAGV 248
 DB 156 GKQGYDDNGHGTHVAGDAAGNGVYASNGYKQVAPBANIVAKVLDADAGRLSSDILGM 215
 QY 249 DWVONKDKYKIRVINLSLSSQSSDG-TDSLQAVNNAMPAGIVVCAAGNSGPNYTV 307
 DB 216 QWILNKKXKYNIRVLSIGETPSLPAPLPLVRSVDLTMNGLVVTVAAAGSGENYTI 275
 QY 308 GSPAAASKVITTYGAND-----SNDNIASFSSRGPTADGRLKEPVAVPGVDIIAPRASG 360
 DB 276 TSPGTSKQALIVGAVDDKRTSDISDETAQPSRG-SPLYKPDIVAPGVKIVSTASEN 333
 QY 361 TSMG---TPINDYTKAGTSMATPHVSGVALLIOAPSPWTPDKVKTALLETADIVAPK 417
 DB 334 VPFADDEVITINKAYATATGTSMATPHMAGAAALLERKNPLTNQIKILASTA--IKID 391
 QY 418 EADIYAGAGRVYVYKAIK 436
 DB 392 DAGLMTQSGSMINTEALK 410

RESULT 13
 O82B14 STRAM PRELIMINARY; PRT; 1208 AA.
 AC O82B14;
 DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
 DT 01-JUN-2003, sequence version 1.
 DE 07-FEB-2006, entry version 19.
 DE Putative protease.
 GN OrderedLocustNames=SAV5721;
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomyces.
 NC NCBI_TaxID=33903;
 RN [1]

Db 293 FEMAVNN---NARIISFGSGSPEH---DSLFTTMINKVVAAGVPIVIAAGNDGSGGT 344
 QY 307 VCSPPAASKVITYGAVDSNDINIASFSRGP-TADGR--LKEPVAPGVDIAPRAGTSM 363
 Db 345 ITPGGEINSLTIGADDSADLADFSRGPVITDDQYIRPDLSAREVSI-----P 395
 QY 364 GRPINDYTKASGTSMATPHVSGVALLIOAHPSWTPDKVKTALILETA-DIVAPKEIADI 422
 Db 396 STVPGGAYAGDGTSMAPHSVSGTVALILEKPTMTPEAVKKELESTAVDLGSAGK--DN 453
 QY 423 AYAGAGVNVYKAL-----KDDVYAKLTFPGSVADKSGATTFPVSGATFYATLTYMDTGS 477
 Db 454 DYSSGRDAVDAVFGKGPVAFYAKPT-SGKVPVLVAFDT-----STGPTTKRMNFGD 507
 QY 478 SDIDLVLYPDNGBVDVSYTAAVYGFEXVYNNPAGTMTVK--VVSYKG-----AANYQV 530
 Db 508 GS-KSFLQNPK-----HKYS-----KAGTYTNLVKAKAGKNTYTKTDYIV 548
 QY 531 DVVSDGSLSGSGGNPNPNPNPTPTDQTFGTGSVN-DYMDTSDTFMNVNSGATKIT 569
 Db 549 -VISKPTAFRSA-----SPTSGKAPLTVAFDTKSSGDNPTAMKWS-----F 587
 QY 590 GDITPTSTYVDLILYDPNGNL---VDRSTSSNSYEHVZY---ANPARQWT 636
 Db 588 GDGTISREKNPTRQYLOEGKYKITLTVSNAGSKTTKINYIKVTTTRBGIYS 641

RESULT 15

OSJ315_GEOKA PRELIMINARY; PRT; 442 AA.
 ID OSJ315_GEOKA
 AC OSJ315;
 DT 01-FEB-2005, integrated into UniProtKB/TrEMBL.
 DT 01-FEB-2005, sequence version 1.
 DT 07-FEB-2006, entry version 10.
 DE Intracellular alkaline serine protease.
 GN OrderedLocustNames=GK0210;
 OS Geobacillus kaustophilus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
 OX NCBI_TaxID=1462;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=HTA426;
 RX PubMed=15576355; DOI=10.1093/nar/gkh970;
 RA Takami H., Takaki Y., Chee G.-J., Nishi S., Shimamura S., Suzuki H., Matsui S., Uchiyama I.;
 RA "Thermoadaptation trait revealed by the genome sequence of thermophilic Geobacillus kaustophilus.";
 RT thermophilic Geobacillus kaustophilus.";
 RL Nucleic Acids Res. 32:6292-6303(2004).
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivative license
 CC
 DR EMBL, BA000043; BAD74495.1; -, Genomic DNA.
 DR GO; GO:0008233; F:Dependase activity; IEA.
 DR GO; GO:0042802; F:Protein self binding; IEA.
 DR GO; GO:004289; F:Subtilase activity; IEA.
 DR GO; GO:0043086; P:negative regulation of enzyme activity; IEA.
 DR GO; GO:0006508; P:proteolysis; IEA.
 DR InterPro; IPR000209; Pept_S8_S53.
 DR InterPro; IPR010259; Prot_inh_S8A.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF05922; Subtilisin_N; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN_1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KM Complete proteome; Hydrolase; Protease; Serine protease.
 SQ SEQUENCE 442 AA; 48629 MW; 30F9F760AD09D20 CRC64;

Query Match 18.9%; Score 649.5; DB 2; Length 442;
 Best Local Similarity 40.6%; Pred. No. 4.3e-27;
 Matches 152; Conservative 65; Mismatches 120; Indels 37; Gaps 11;

QY 85 QVKSYSKIIPAAVAKI-KARDLLIAGMIDTGYFGNTRVSGIKFIOEDYKVOVDATSV 142
 Db 79 KYVHHFRHVPFYSARVTPALAEQLLEHPRKYKVFENRT-----VKALLNNAVPS 127
 QY 143 SQIGADTWNNSLGYDSGVVVAIVDNGIDANHPDLKGKVLGWDAVNGRSTPYDDOCHGT 202
 Db 128 ANAKRVAV-NGTELSKGVTLIVDTGI-YPHPDLEGRIAAFVDFVNGKRTTPYDDNGHGT 185
 QY 203 HVAGIVAGTGSV-NSOYIGVAPGAKLVGVKVLGADSGSVSTIIAGVDVWVQNKDKY--- 258
 Db 186 HCAGDAAAGRGMDGDIYAGPAVEANLIGVKVLDRSSGSGLETIMRGIEWCIDVNEHPSK 245
 QY 259 GIRVINLSLSSGSS--DGTSLSDAVNNAMDAGIVCYAAGNSGPNYTTVGSPPAAAK 315
 Db 246 RIDIILSLSGEPQPPPIENDDELVQVAAQAMEQGIIVCAAGNEGPNYGTISSPGISDR 305
 QY 316 VITVGAVD-----SNDINIASFSRGPAPDRLEPEVAPGVDI---APRASGTSMG 364
 Db 306 IITVGALEDHDTATTRADDVASFSSRGPTEYGVTPDLVPEVNILSLRAPSFLDKKN 365
 QY 365 --TPINDYTKASGTSMATPHVSGVALLIOAHPSWTPDKVKTALILETADIAPKEIADI 422
 Db 366 KQSRVGDHIIISGTSMTATPICAGIYALMLQAKENATPDEIKRALDGDADLMKGRD--PV 423
 QY 423 AYAGAGVNVYKAIK 436
 Db 424 VYAGGYVNGRAIE 437

Search completed: January 6, 2007, 22:31:24
 Job time: 194 secs

November 2005

Published Applications Nucleic Acid and Published Applications Amino Acid database searches now generate two sets of results each. The Published Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published Applications New databases; older published applications make up the Published Applications Main databases.

Searches run against Nucleic Acid Published Applications produce two sets of results, with the extensions **.rnpbm** (Published Applications NA Main) and **.rnpbn** (Published Applications NA New).

Searches run against Amino Acid Published Applications produce two sets of results, with the extensions **.rapbm** (Published Applications AA Main) and **.rapbn** (Published Applications AA New).

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2007 Bioceleration Ltd.

OM protein - protein search, using SW model

Run on: January 6, 2007, 22:33:57 ; Search time 31 Seconds
(without alignments)
2087.781 Million cell updates/sec

Title: US-10-800-684-1
Perfect score: 3437
Sequence: 1 MKRLGAVTLALVGLAGT.....YAVSTGYMADYQKAVVYVG 659

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 393712 seqs, 98211237 residues
Total number of hits satisfying chosen parameters: 393712

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Published Applications AA New:
1: /EMC Celerra_SIDS3/Ptodata/1/pubpaa/US09_NEM_PUB.pep.*
2: /EMC Celerra_SIDS3/Ptodata/1/pubpaa/US06_NEM_PUB.pep.*
3: /EMC Celerra_SIDS3/Ptodata/1/pubpaa/US07_NEM_PUB.pep.*
4: /EMC Celerra_SIDS3/Ptodata/1/pubpaa/US08_NEM_PUB.pep.*
5: /EMC Celerra_SIDS3/Ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
6: /EMC Celerra_SIDS3/Ptodata/1/pubpaa/US10_NEM_PUB.pep.*
7: /EMC Celerra_SIDS3/Ptodata/1/pubpaa/US11_NEM_PUB.pep.*
8: /EMC Celerra_SIDS3/Ptodata/1/pubpaa/US60_NEM_PUB.pep.*

* Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 2664 | 77.5 | 663 | 6 | US-10-526-324-251 |
| 2 | 632.5 | 18.4 | 443 | 6 | US-10-530-643-8 |
| 3 | 616.5 | 17.9 | 1227 | 6 | US-10-530-643-6 |
| 4 | 576.5 | 16.8 | 552 | 6 | US-10-530-643-124 |
| 5 | 559.5 | 16.3 | 379 | 7 | US-11-433-614-14 |
| 6 | 546.5 | 15.9 | 1606 | 6 | US-10-530-643-114 |
| 7 | 523.5 | 15.2 | 275 | 6 | US-10-567-073-3 |
| 8 | 523.5 | 15.2 | 275 | 6 | US-10-541-737-1 |
| 9 | 521.5 | 15.2 | 274 | 6 | US-10-532-605-2 |
| 10 | 521.5 | 15.2 | 274 | 6 | US-10-541-737-3 |
| 11 | 516.5 | 15.0 | 381 | 7 | US-11-452-695-2 |
| 12 | 513.5 | 14.9 | 381 | 7 | US-11-452-695-3 |
| 13 | 513.5 | 14.9 | 381 | 7 | US-11-452-695-4 |
| 14 | 508.5 | 14.8 | 524 | 6 | US-10-526-324-20 |
| 15 | 507.5 | 14.6 | 381 | 7 | US-11-433-614-16 |
| 16 | 500.5 | 14.6 | 275 | 6 | US-10-541-737-2 |
| 17 | 500 | 14.5 | 624 | 6 | US-10-530-643-46 |
| 18 | 490.5 | 14.3 | 673 | 6 | US-10-530-643-56 |
| 19 | 488 | 14.2 | 462 | 6 | US-10-449-902-40515 |
| 20 | 487.5 | 14.2 | 269 | 6 | US-10-541-737-4 |
| 21 | 486 | 14.1 | 274 | 7 | US-11-452-695-1 |
| 22 | 484 | 14.1 | 400 | 6 | US-10-530-643-60 |
| 23 | 476.5 | 13.9 | 518 | 6 | US-10-530-643-122 |
| 24 | 476 | 13.8 | 422 | 6 | US-10-526-324-979 |
| 25 | 470.5 | 13.7 | 469 | 6 | US-10-449-902-37064 |

| | | | | | | |
|----|-------|------|-----|---|-------------------|--------------------|
| 26 | 469 | 13.6 | 640 | 7 | US-11-318-576-4 | Sequence 4, Appl1 |
| 27 | 468.5 | 13.6 | 523 | 6 | US-10-530-643-4 | Sequence 4, Appl1 |
| 28 | 465 | 13.5 | 607 | 6 | US-10-530-643-62 | Sequence 62, Appl1 |
| 29 | 457.5 | 13.3 | 434 | 7 | US-11-318-576-10 | Sequence 10, Appl1 |
| 30 | 454.5 | 13.2 | 615 | 6 | US-10-530-643-88 | Sequence 88, Appl1 |
| 31 | 452.5 | 13.2 | 434 | 7 | US-11-318-576-2 | Sequence 2, Appl1 |
| 32 | 452.5 | 13.2 | 434 | 7 | US-11-318-576-11 | Sequence 11, Appl1 |
| 33 | 450.5 | 13.1 | 434 | 7 | US-11-318-576-15 | Sequence 15, Appl1 |
| 34 | 443 | 12.9 | 433 | 7 | US-11-318-576-12 | Sequence 12, Appl1 |
| 35 | 443 | 12.9 | 433 | 7 | US-11-318-576-14 | Sequence 14, Appl1 |
| 36 | 438 | 12.7 | 433 | 7 | US-11-318-576-13 | Sequence 13, Appl1 |
| 37 | 433.5 | 12.6 | 269 | 6 | US-10-516-164A-1 | Sequence 1, Appl1 |
| 38 | 432.5 | 12.6 | 433 | 7 | US-11-318-576-16 | Sequence 16, Appl1 |
| 39 | 432.5 | 12.6 | 611 | 6 | US-10-530-643-74 | Sequence 74, Appl1 |
| 40 | 429 | 12.5 | 434 | 6 | US-10-530-643-2 | Sequence 2, Appl1 |
| 41 | 423 | 12.3 | 426 | 6 | US-10-530-643-42 | Sequence 42, Appl1 |
| 42 | 413 | 12.0 | 412 | 6 | US-10-530-643-128 | Sequence 128, App |
| 43 | 407 | 11.8 | 515 | 6 | US-10-530-643-50 | Sequence 50, Appl1 |
| 44 | 397.5 | 11.6 | 579 | 6 | US-10-530-643-66 | Sequence 66, Appl1 |
| 45 | 396.5 | 11.5 | 579 | 6 | US-10-530-643-70 | Sequence 70, Appl1 |

ALIGNMENTS

RESULT 1
US-10-526-324-251
Sequence 251, Application US/10526324
Publication No. US20060248617A1
GENERAL INFORMATION:
APPLICANT: Imanaka, Takayuki
TITLE OF INVENTION: METHOD OF TARGETED GENE DISRUPTION, GENOME OF
TITLE OF INVENTION: HYPERTHERMOSTABLE BACTERIUM AND GENOME CHIP USING
FILE REFERENCE: 490051.401USPC
CURRENT FILING DATE: 2005-02-28
PRIOR APPLICATION NUMBER: US/10/526,324
PRIOR FILING DATE: 2003-08-29
PRIOR APPLICATION NUMBER: PCT/IB2003/003597
PRIOR FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: JP 2002-319011
NUMBER OF SEQ ID NOS: 2167
SOFTWARE: PatentIn version 3.1
SEQ ID NO 251
LENGTH: 663
TYPE: PRT
ORGANISM: Thermococcus kodakarensis KOD1
FEATURE:
NAME/KEY: misc feature
LOCATION: (414562)..(414542)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc feature
LOCATION: (786890)..(786890)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc feature
LOCATION: (786907)..(786907)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc feature
LOCATION: (786944)..(786946)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc feature
LOCATION: (839139)..(839139)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc feature
LOCATION: (1128488)..(1128488)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:

```

NAME/KEY: misc feature
LOCATION: (1128499)..(1128499)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc feature
LOCATION: (1128505)..(1128506)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc feature
LOCATION: (1128517)..(1128518)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc feature
LOCATION: (1128539)..(1128540)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc feature
LOCATION: (1561400)..(1561400)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc feature
LOCATION: (1561477)..(1561477)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc feature
LOCATION: (1561545)..(1561545)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc feature
LOCATION: (1767941)..(1767941)
OTHER INFORMATION: n is a or c or g or t.
US-10-526-324-251

```

```

Query Match 77.5%; Score 2664; DB 6; Length 663;
Best Local Similarity 75.9%; Pred. No. 1,3e-181;
Matches 505; Conservative 64; Mismatches 88; Indels 8; Gaps 4;

```

```

QY 1 MKRGAVVLAALVGLAGTALPAVPEVVRNNAVQCKNTGLTPGLFKYQVRNNQEV 60
DB 1 MKRGAVVLAALVGLAGTALPAVPEVVRNNAVQCKNTGLTPGLFKYQVRNNQEV 59
QY 61 DTVMFSGYGRDRRAVKRLRMGAOVKYSYKIPAVVVKIKAPDLIIAMIDTGYPGNT 120
DB 60 STIMFNOADKEKAVILDFLAKIKYNNHIIIPALVAKIKVLDLIIAGLMDTGFFGNA 119
QY 121 RVSGIKRIGEDYKQVDDAT----SVSOIGADVTWNSLGYDSGVVAIVDTGIDANHP 175
DB 120 QLSGVCFIOEDYVVKVAVETEGEDESAQVMAITMMN-LOYDSGIIIGIIDGIDASHP 178
QY 176 DLGKVIQWDAVNGSTPYDDOGHGHVAGIYAGTGSV-NSOYIGVAPAKIVGYKVLG 234
DB 179 DLGKVIQWDAVNGSTPYDDOGHGHVAGIYAGTGSV-NSOYIGVAPAKIVGYKVLG 234
QY 235 ADSSGSVTIIAGVWVONKDKYKIRVIMLSGSSOSSDGTDSLQAVNANMADAGIYVC 244
DB 239 GQSSGSISDIINGVDAVONKDKYKIRVIMLSGSSOSSDGTDSLQAVNANMADAGIYVC 298
QY 295 VAAAGSGPNTYVGSPPAASKVITVGAVDSNDNIASFSSRGPADGRLKPEVVAPEGVDII 354
DB 299 VAAAGSGPNTYVGSPPAASKVITVGAVDSNDNIASFSSRGPADGRLKPEVVAPEGVDII 358
QY 355 APAASGSMGTPIINDYTKASGTSMATPHVSGVALLIOAHPSWTDPKVTALIEADIV 414
DB 359 APAASGSMGTPIINDYTKASGTSMATPHVSGVALLIOAHPSWTDPKVTALIEADIV 418
QY 415 APKEIADIAAGRVNYYKAIKTDYAKLFTGSAVDKSGATTFPVSGATFTATLYMD 474
DB 419 KPBEIADIAAGRVNYYKAIKTDYAKLFTGSAVDKSGATTFPVSGATFTATLYMD 478
QY 475 TGSSDIDLILYDNGNEVDYSYTAAYGFEKVGYNPAGTWTVKVSYKGAANYQVDVVS 534
DB 479 NSGSDIDLILYDNGNEVDYSYTAAYGFEKVGYNPAGTWTVKVSYKGAANYQVDVVS 538

```

```

QY 535 DGSLSGSGGNPNPNPNTPTTDPOTFNQSVNDYDSDTFETMNVNSGATKIGDTLP 594
DB 539 DGSLSGSGSGGSEPSSEPPYDEKTFGVHDYDSDTFETMNVNSGATKIGDTLP 598
QY 595 DTSYNDIDLILYDNGNLVDRSTSSNSYEHVEYANPAPGTWFLVYAYSTYGADYOLKA 654
DB 599 DTSYNDIDLILYDNGNLVDRSTSSNSYEHVEYANPAPGTWFLVYAYSTYGADYOLKA 658
QY 655 VVYVG 659
DB 659 KVVYVG 663

```

RESULT 2

```

US-10-530-643-8
Sequence 8, Application US/10530643
Publication No. US2006025995A1
GENERAL INFORMATION:
APPLICANT: CAYOUEITE, Michelle
APPLICANT: HANSEN, Connie Jo
APPLICANT: MCCLURE, Amy
APPLICANT: SUN, May
APPLICANT: GRAMATIKOVA, Svetlana
APPLICANT: DYCAICO, Mark
APPLICANT: BARTON, Nelson R.
APPLICANT: STEGE, Justin T.
APPLICANT: ABOUSHADI, Nabla M.
TITLE OF INVENTION: PROTEASES, NUCLEIC ACIDS ENCODING THEM
FILE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
FILE REFERENCE: 564462004100
CURRENT APPLICATION NUMBER: US/10/530,643
PRIOR FILING DATE: 2006-04-07
PRIOR APPLICATION NUMBER: PCT/US03/32819
PRIOR FILING DATE: 2003-10-10
PRIOR APPLICATION NUMBER: 60/471,423
PRIOR FILING DATE: 2003-05-16
PRIOR APPLICATION NUMBER: 60/418,467
PRIOR FILING DATE: 2002-10-10
NUMBER OF SEQ ID NOS: 255
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 443
TYPE: PRT
ORGANISM: Bacteria
US-10-530-643-8

```

```

Query Match 18.4%; Score 632.5; DB 6; Length 443;
Best Local Similarity 36.8%; Pred. No. 3,6e-37;
Matches 153; Conservative 79; Mismatches 143; Indels 41; Gaps 12;

```

```

QY 47 LFKYQVRNNQEV-DEYDVTVMFSGYGRDRRAVKV---LRLMGAOVKYSYKIPAVAVKIK 101
DB 39 LHKMERFRPKKXKKSIVLIEFESNCYQSGCLEVNOAFIRKNCCKIRNEFSIISCSADIT 98
QY 102 ABDLLIACMIDTYFGNTRVSGIKFIOEDYKQV-VDDATSQIGADVTWNSLGYDSG 160
DB 99 PSLLEBVL-----TNCNHIKVVYLNEHVRALLDTAVISANAKIIVRNNTLTGKG 148
QY 161 VVVAIVDGDIDANHPDLGKVIQWDAVNGSTPYDDOGHGHVAGIYAGTGSVNS-QYI 219
DB 149 ITTAVIYDGI-PPHTDLSGRILDPFDINDKREYDDNGHGHGDLGQSSASSGLYM 207
QY 220 GVAAPAKLVGVKVGADSGSVSTIIAGVWV---ONKDYKIRVIMLSGSS--SQS 272
DB 208 GPAPFANVILYGVKLVKVSGLSETIMQGVDMCKIKNDQNPQ-KINIIMLSGAPAQRYE 266
QY 273 SGTGTSLSQAVNANMADAGIYVCVAAAGSGPNTYVGSPPAASKVITVGAVD-----S 324
DB 267 NENDPMTVMYKEMKAWENGIYVCVAAAGSGPASTIASGVEOVITVGALDKTTADTRS 326
QY 325 NDNIASFSSRGPBTADGRLKPEVVAPEGVDIIAPRASGTM-----GTPINDYTKASGTM 379
DB 327 DDEVASFSSRGPBTADGRLKPEVVAPEGVDIIAPRASGTM-----GTPINDYTKASGTM 386

```

QY 380 ATPHVGVALILQAHPSWTPDKVTALJETADIVAPKEIADIAYGARVNYKAI 435
DB 387 ATPICAGIALILIQHNPANTPOEVKOLKRGTDLMTNRD--PNTYAGIYMENSV 440

RESULT 3

US-10-530-643-6
; Sequence 6, Application US/10530643
; Publication No. US2006025995A1
; GENERAL INFORMATION:
; APPLICANT: CAYOJETTE, Michelle
; APPLICANT: CAYOJETTE, Michelle
; APPLICANT: HANSEN, Connie Jo
; APPLICANT: MCCURE, Amy
; APPLICANT: SUN, May
; APPLICANT: GRAMATIKOVA, Svetlana
; APPLICANT: DYCAICO, Mark
; APPLICANT: BARTON, Nelson R.
; APPLICANT: STEGE, Justin T.
; APPLICANT: ABOUSHADI, Nahla M.
; TITLE OF INVENTION: PROTEASES, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564462004100
; CURRENT FILING DATE: 2006-04-07
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: PCT/US03/32819
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/471,423
; PRIOR FILING DATE: 2003-05-16
; PRIOR APPLICATION NUMBER: 60/418,467
; PRIOR FILING DATE: 2002-10-10
; NUMBER OF SEQ ID NOS: 255
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1227
; TYPE: PRT
; ORGANISM: Bacteria
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(32)
US-10-530-643-6

Query Match 17.9%; Score 615; DB 6; Length 1227;

Best Local Similarity 43.3%; Pred. No. 2.6e-35; Indels 26; Gaps 13;
Matches 155; Conservative 55; Mismatches 122;

QY 131 DYKVVQVDDATSVSOLGADTVWNSLGYDSSGVVAIVDTGIDANHPDLKGVIGWYDAVNG 190
DB 192 DGRVKAAMESSNQIETPRKMEK-GILTKGVKVAIVDTGADLHPLDLGRVTESSKFIAG 250
QY 191 RSTPYDDQGHGHTVAGTGS-VNSQYIGVAPKALVGVKVLGADSGSVSTIIAGVD 249
DB 251 OEVA-DRNHGHTVASTVSGSGAGSDGKEKGVAPGATLAVGKVLSDGSGSESEIIAGME 309
QY 250 WVVQVQNDKXGIRVINSLSGSSSDGTDLSQAVNN-AMDAIGIVCVAAGNSG-PNTYTY 307
DB 310 WAAKDD---AKVSNLSGSRPSDGTDPVALVNLTLTETGALFVYLAGNSGPG--SI 364
QY 308 GSEPAASKYITVAVVNSNDNIAFSRSRGPT-ADGRLEKPEVAVGVDIIPRAGSGTSGTP 366
DB 365 GSGPAGDASLITIGAVVDSABEAFTSQGPRYGDQALKPDISAPGVDIILAR-----SGLLP 420
QY 367 INDYTKASGTSNATPHVSGVALILQAHPSWTPDKVTALJETADIVAPKEIADIAYG 426
DB 421 GSGLYTSMGTSNATPHVAGVALLAERHPDWTGADLKCALMSSSKTL--DASSYALGS 477
QY 427 GRNVVVKAIKYDDYAKLTFPTGSVADKGSATHTFDVSGATVTLNLTWDGSSPIDLYL 484
DB 478 GRVDVAAALIA---ANVTATGS-ADIGFVAMPYASSKPYTKVTVY--TNSSDAPVEL 527

RESULT 4
US-10-530-643-124

; Sequence 124, Application US/10530643
; Publication No. US2006025995A1
; GENERAL INFORMATION:
; APPLICANT: CAYOJETTE, Michelle
; APPLICANT: HANSEN, Connie Jo
; APPLICANT: MCCURE, Amy
; APPLICANT: SUN, May
; APPLICANT: GRAMATIKOVA, Svetlana
; APPLICANT: DYCAICO, Mark
; APPLICANT: BARTON, Nelson R.
; APPLICANT: STEGE, Justin T.
; APPLICANT: ABOUSHADI, Nahla M.
; TITLE OF INVENTION: PROTEASES, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564462004100
; CURRENT FILING DATE: 2006-04-07
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: PCT/US03/32819
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/471,423
; PRIOR FILING DATE: 2003-05-16
; PRIOR APPLICATION NUMBER: 60/418,467
; PRIOR FILING DATE: 2002-10-10
; NUMBER OF SEQ ID NOS: 255
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 124
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample.
; NAME/KEY: SIGNAL
; LOCATION: (1)...(56)
US-10-530-643-124

Query Match 16.8%; Score 576.5; DB 6; Length 552;
Best Local Similarity 43.7%; Pred. No. 4.8e-33; Indels 61; Gaps 9;
Matches 141; Conservative 37; Mismatches 84;

QY 158 GSGVVAIVDTGIDANHPDLK-----KYIGWYDAVNGRSTPYDDQGHGHTVA 205
DB 245 GSGVTAIITDGTVDYHNPDLGGCTTQEFLGACEKVIYGVDFINDDAMDNDGHTVA 304
QY 206 GIVAGTGSVNSQYIGVAPKALVGVKVLGADSGSVSTIIAGVDVWVQNDKXGIRVIN 255
DB 305 GIAAGGGLK---GVAAPAKILAYVGLAGGVGTWEGIIAGIBoVID---GADILSL 356
QY 266 SIGSSGSDGTDLSL-SQAVNNAMDAIGIVCVAAGNSGPTTYTGVSPAAASKYITVGA VDS 324
DB 357 SIGCHSSCNPDIDIASQAVDNVLAQKVVVVAAGNSGSSRTIGSGPTARKAITVGS TTK 416
QY 325 NDNIAFSRSRGPTA-----DGRLEKPEVAVGVDIIPRAGSGTSGTP----- 366
DB 417 SIIISFSSRGFVMMDBAGIEQAIKRPVLA PG-----GTDGSEFNCPEMMFDN 467
QY 367 -----INDYTKASGTSNATPHVSGVALILQAHPSWTPDKVTALJETADIVA--PKG 418
DB 468 RICAMWLNEKYLAISGTSNATPLVSGAIALNOKHPDWTPEELKGAIVKGTAINLGYDPNE 527
QY 419 IADIAVAGRVVYVYKIKYDDVA 441
DB 528 -----QAGRINVRMIGLEBRA 545

RESULT 5
US-11-433-614-14
; Sequence 14, Application US/11433614
; Publication No. US20060205019A1
; GENERAL INFORMATION:
; APPLICANT: Athena Biotechnologies, Inc.
; APPLICANT: Menzel, Rolf
; TITLE OF INVENTION: Method and Compositions for Directed Gene Assembly

```

; FILE REFERENCE: 46675-5004-01-US
; CURRENT APPLICATION NUMBER: US/11/433,614
; CURRENT FILING DATE: 2006-05-12
; PRIOR APPLICATION NUMBER: US 09/920,118
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent version 3.3
; SEQ ID NO 14
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-11-433-614-14

```

```

Query Match      16.3%; Score 559.5; DB 7; Length 379;
Best Local Similarity 38.2%; Pred. No. 4.6e-32;
Matches 147; Conservative 54; Mismatches 135; Indels 49; Gaps 13;

```

```

QY 57 NQEVDTVMFGSYGDRDRAVK--VLRLMGAQVKSXKTIIPAVAVKIKARDLLIAGMIDT 114
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 35 NVEKDYIVGFKS-GVKTKASVKKDIIKSGGKVDKQFRILINAKAKLDEALEEVKNDPDV 93
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 115 GFQNTRVSGIKFIOEDYKVVDDAT---SVSOIGADTVWNSLGVDGSGVVVAIVDTGID 171
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 94 AY-----VEEDHVAHALAQVTPYGIPIKADKV-OAQGYKGAIVKAVAVLDTGIQ 141
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 172 ANHPDLKGVIGWYDAVNGRSTPYDDQHGHTHVAIGVAGSVNSOYIGVAPGAKLVGVK 231
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 142 ASHDDL--NVVGGASFYVAGEAYNTDGNHGHVAGTVAALDNTTG-VLGVAPVNSLVAVK 198
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 232 VLGADGSGSVSTIIAGVDMVYQNKDKGIRVINSLSGSSQSDGTSLSQAVNNADAGT 291
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 199 VINSGSGSVSGISGISEMATTN---GMDVINMSIG--GPSGSTMKAQVNAAYVARGV 251
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 292 VVCAAGNSGP--NTYTVGSPAAASKYITVGAVDNSNINASFSSRGPTADGRLEKPEVAP 349
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 252 VVVAAGNSGSGSNTNTIGYPAKYDSIVAVGADVDPNSNRASFSSVG-----AELEVAP 305
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 350 GVDIAPRASGTSMTGTINDYTKASGTSMTATPHVSGVALIIQAPSWTPDKVKTALIE 409
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 306 GAGVYSTYPTST-----YATLNGTSMASPHVAGAAALILSKHPNLSASGVRRRLSG 356
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 410 TADIVAPKEIADIVAGGAVNVYKA 434
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 357 TATYLG---SPTYGKGLINVEAA 377
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

RESULT 6

US-10-530-643-114

Sequence 114, Application US/10530643

Publication No. US2006025995A1

GENERAL INFORMATION:

APPLICANT: CAROUEITE, Michelle

APPLICANT: HANSEN, Connie Jo

APPLICANT: MCCLURE, Amy

APPLICANT: SUN, May

APPLICANT: GRAMATIKOVA, Svetlana

APPLICANT: DYCATCO, Mark

APPLICANT: BARTON, Nelson R.

APPLICANT: STEGE, Justin T.

APPLICANT: ABUSHADI, Nahla M.

TITLE OF INVENTION: PROTEASIS, NUCLEIC ACIDS ENCODING THEM

FILE REFERENCE: 564462004100

CURRENT APPLICATION NUMBER: US/10/530,643

CURRENT FILING DATE: 2006-04-07

PRIOR APPLICATION NUMBER: PCT/US03/32819

PRIOR FILING DATE: 2003-10-10

PRIOR APPLICATION NUMBER: 60/471,423

PRIOR FILING DATE: 2003-05-16

PRIOR APPLICATION NUMBER: 60/418,467

PRIOR FILING DATE: 2002-10-10

NUMBER OF SEQ ID NOS: 255

SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 114
; LENGTH: 1606
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample.
; NAME/KEY: SIGNAL
; LOCATION: (1)...(32)
US-10-530-643-114

```

```

Query Match      15.9%; Score 546.5; DB 6; Length 1606;
Best Local Similarity 23.4%; Pred. No. 2.9e-30;
Matches 240; Conservative 110; Mismatches 222; Indels 455; Gaps 41;

```

```

QY 7 VVALVVLVGLAAGTALAAPYKPVVNN---AVQK---NYGLLTPG--LFFKVGQM 54
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 16 LINVLSLVLPFNGTANAALLKQPFKTPQDASALQOKAIAQOTSLLGEDARLHKDQL 75
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 55 NMQEVDVIMFGSYGDRBRAV---KVLRLMGAQVK----- 87
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 76 SSGQEVPIIHS---EKAVGLEQGIHKLKSKMSQSEITLKKKGTQAQONTAKEMT 130
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 88 -----YSY-KIIPAVAVKIKARDLLIAGMIDTGYFGNTRVSGIKFIOEDYKV--- 134
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 131 YKKVKFPDYVYDVTVLNGLSGTVOANDLKL-----LITSGVKVPEPDTTVAQ 179
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 135 ----QVD---DA---TSVSOIGADTVWNSLGVDGSGVVVAIVDTGIDANPDLK-----G 179
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 180 EGLQKVPYVDKMDTSLPFLGINKLMEB-GIEGQGVKVAIVDTGIDASHPDLKAVYKGG 238
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 180 K-----VIGWY-----DANGRSTPD-----DQ-----HGTVAIVAGTGSV 214
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 239 KMFVHVDSFYARPRADDGSETSPIDRANKPEVNERSSYTHGHVAGTIVAQGN 298
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 215 NSQYIGVAPGAKLVGVKVLGADGSGSVSTIIAGVDMVYQNKDKYGIRVINSLSGSSQSD 274
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 299 EFGIKGIAPKVDLYSVRYLVGAVGSGATSGI IKAIDTAVIEK---IDVINSLGGANSE 354
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 275 GTDSLSQAVNNAMDGIYVCVAAAGNSGRTTYVSGPAAASKYITVG----- 320
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 355 -TDGASFALNNMMLAGTISVATGNSGPRGTGTPATSRIGIAVGNTPBETHQAVNN 413
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 321 ----- 320
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 414 VKAGEFTSKVNDLMATTFGKQDETQLAGYDLIAPGAGAKYADLDVTGKALVARS 473
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 321 ---AVD----- 323
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 474 EIAFVDKVAALKONGAVAIIYHNFAGTNAAPNKSDFLGDAAFEFIPFDMGVTDOEALRA 533
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 324 -----SNDNIAFSSRGP-TADGSLKPEVAVPGVDIIA--PRASGT 361
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 534 ALAAAPSTISPDQSSIKTTGDVDNDSSRGSPFNFIDIKPVSAFGTNISSTIPMYGND 593
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 362 SMGTPIINDYTKASGTSMTATPHVSGVALIIQAPSWTPDKVKTALITADIVAPKEIAD 421
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 594 PPDADYSEAYSRKTSMTATPHIAGIALVQGANPTMVAFPVKYALSTATAVLDTKKIDV 653
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 422 IAYGAGRVNTYKAYD----- 438
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 654 PAQAGRYDAVKARADVLAYAIPTASNDGTEVENLKGTVPFGPKLDKNISYTKIKNVK 713
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 439 -----DYA-----KLTPGVSAD----- 451
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 714 DKSAAGDYTVSVDTYGFQDPAKTYVDQSEFTLNGEQLLVTLTASAKETAGDELIGYI 773
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 452 --KG-----SATHTFDVSGATFVT-----ATLYW-----DTGSSD 479
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 774 HIKGNGKDLSPFAADGGAATAVAVKMEIKTIDLSFNGSGVNDADALYFIITSDVGFNT 833
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 480 IDLY-LIDPNKNEVDYSTAYYGEKVGYYNP---TAGTWTV---KVVSYKGAANYOV 530
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

Db 834 IBLMDIMNPTGKGYSDGY-----IGYLHSDTLTLAGSYCLPLIKYTOGNGTAS--- 882
QY 531 DVSODSL-----SQSGGNPNPNPN-PPPTPTDQ--PFTGSVND-----YMD- 572
Db 883 EVIPLDLYTIDPSAETKSNPTIADYVGPVVVKSTAGTIEGAVADGKVTGKTIDKVDY 942
QY 573 -----TSDPTMNVN-----SGATKITGD-----LTFDTSYNDLDLY 604
Db 943 QKELVKGWGYDLNLTSLATYEVTANADVDSGAVLADGTFAPLPTFDKTKNNVTYK 1002
QY 605 LYDPNGN 611
Db 1003 YADAAGN 1009

RESULT 7
US-10-567-073-3
; Sequence 3, Application US/10567073
; Publication No. US20060134740A1
; GENERAL INFORMATION:
; APPLICANT: Bryan, Philip N.
; TITLE OF INVENTION: Engineered Proteases for Affinity Purification and Processing of
; TITLE OF INVENTION: Fusion Proteins
; FILE REFERENCE: 4115-181
; CURRENT APPLICATION NUMBER: US/10/567, 073
; CURRENT FILING DATE: 2006-02-03
; PRIOR APPLICATION NUMBER: US 60/493,032
; PRIOR FILING DATE: 2003-08-06
; PRIOR APPLICATION NUMBER: PCT/US04/021049
; PRIOR FILING DATE: 2004-06-29
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Bacillus amyloliquefaciens
US-10-567-073-3

Query Match 15.2%; Score 523.5; DB 6; Length 275;
Best Local Similarity 44.4%; Pred. No. 1.1e-29;
Matches 132; Conservative 40; Mismatches 90; Indels 35; Gaps 11;

QY 142 VSQIGADTVWNSLGYDGSVVAIVDTGIDANHPDLKGVIGWYDAVNGRSTPY--DDQGH 200
Db 8 VSQIKAPAL-HSQGYTGSNVKVAVIDSGIDSSHPLD--KVAGASVNPSETNPFQDNNSH 64
QY 201 GTHVAGIVAGTGSVNSQYIGVAPGAKLVGVKVLGADGSGSVSTIIAGVDWVQNKDKYGI 260
Db 65 GTHVAGTVAALNN-SIGVLGVAPASLVAVKVLGADGSGQYMWIINGIEMALAN---NM 119
QY 261 RVINISLSSGSSDGTDSLQAVNNAMDAGIYVCVAAGNSGP--NTYTVGSPAAASKVIT 318
Db 120 DVINNLSLG---GPSGSAAKAAVDKAVAGVVAAGNMGTSSTVGPCKYPSVIA 176
QY 319 VGANVNDNIASFSSRGPTADRLKPEVVAAPVDIIAPASGTSMTGPIINDYTTKASGTS 378
Db 177 VGANVSSNORASFSSVGPBLD-----VMAPGVSI-----QSTLPKNTGAYANGTS 221
QY 379 MATPHVSGVALILQAHPSWTPDKVTALIFETADIAPKEIAD-IAYGAGRNVVYKA 434
Db 222 MASPHVAGAAALILSKHPMTWTQVRSLENTT-----TKLGDSPFYGKGLINVQA 273

RESULT 8
US-10-541-737-1
; Sequence 1, Application US/10541737
; Publication No. US20060252155A1
; GENERAL INFORMATION:
; APPLICANT: Leeifang, Chris
; APPLICANT: Van der Kleij, Wilhelmus A.H.
; TITLE OF INVENTION: Methods for Site-Directed Mutagenesis
; TITLE OF INVENTION: and Targeted Randomization
; FILE REFERENCE: GC787-2

; CURRENT APPLICATION NUMBER: US/10/541,737
; CURRENT FILING DATE: 2005-07-08
; PRIOR APPLICATION NUMBER: PCT/US04/01334
; PRIOR FILING DATE: 2004-01-16
; PRIOR APPLICATION NUMBER: US 60/440,792
; PRIOR FILING DATE: 2003-01-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Bacillus amyloliquefaciens
US-10-541-737-1

Query Match 15.2%; Score 523.5; DB 6; Length 275;
Best Local Similarity 44.4%; Pred. No. 1.1e-29;
Matches 132; Conservative 40; Mismatches 90; Indels 35; Gaps 11;

QY 142 VSQIGADTVWNSLGYDGSVVAIVDTGIDANHPDLKGVIGWYDAVNGRSTPY--DDQGH 200
Db 8 VSQIKAPAL-HSQGYTGSNVKVAVIDSGIDSSHPLD--KVAGASVNPSETNPFQDNNSH 64
QY 201 GTHVAGIVAGTGSVNSQYIGVAPGAKLVGVKVLGADGSGSVSTIIAGVDWVQNKDKYGI 260
Db 65 GTHVAGTVAALNN-SIGVLGVAPASLVAVKVLGADGSGQYMWIINGIEMALAN---NM 119
QY 261 RVINISLSSGSSDGTDSLQAVNNAMDAGIYVCVAAGNSGP--NTYTVGSPAAASKVIT 318
Db 120 DVINNLSLG---GPSGSAAKAAVDKAVAGVVAAGNMGTSSTVGPCKYPSVIA 176
QY 319 VGANVNDNIASFSSRGPTADRLKPEVVAAPVDIIAPASGTSMTGPIINDYTTKASGTS 378
Db 177 VGANVSSNORASFSSVGPBLD-----VMAPGVSI-----QSTLPKNTGAYANGTS 221
QY 379 MATPHVSGVALILQAHPSWTPDKVTALIFETADIAPKEIAD-IAYGAGRNVVYKA 434
Db 222 MASPHVAGAAALILSKHPMTWTQVRSLENTT-----TKLGDSPFYGKGLINVQA 273

RESULT 9
US-10-532-605-2
; Sequence 2, Application US/10532605
; Publication No. US20060134092A1
; GENERAL INFORMATION:
; APPLICANT: Miwa, Takenori
; TITLE OF INVENTION: Method for digesting proteins highly resistant to denaturation
; FILE REFERENCE: 087625
; CURRENT APPLICATION NUMBER: US/10/532,605
; CURRENT FILING DATE: 2005-04-25
; PRIOR APPLICATION NUMBER: PCT/JP2003/013658
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-532-605-2

Query Match 15.2%; Score 521.5; DB 6; Length 274;
Best Local Similarity 41.9%; Pred. No. 1.5e-29;
Matches 131; Conservative 39; Mismatches 94; Indels 49; Gaps 10;

QY 124 GIKFIQEDYKVQVDDATSVSQIGADTVWNSLGYDGSVVAIVDTGIDANHPDLKGVIG 183
Db 7 GIPLIKAD-KYQAQ-----CYKANVYVGIIDTGIASSHDL--KVVG 46
QY 184 WYDAVNGSTPYDDOQGHGTHVAGIYAGTGSVNSQYIGVAPGAKLVGVKVLGADGSGSVST 243
Db 47 GASPFSGESYNDGNGHGHVAGTVAALDNTTG-VLGVAPVNSLVAIKVLNAGSGGYVSA 105
QY 244 ILAGVDWVQNKDKYGIKVINISLSSGSSDGTDSLQAVNNAMDAGIYVCVAAGNSGP- 302

```
Db 106 IYSGIMATON-----GIDVIMSLG---GPSGSTALKQAVDKAVASGIVVAAAGSGSGS 158
QY 303 -NNTYTGSPAAASKVITVGVAVDSNDNIASFSSRGPTADGRLKPEVVAAPGVDIAPRASGT 361
Db 159 GSQNTTGYPKYDSVIAVGAVDNSKNRASFSSVGS-----ELEVAAPGVSVSTYPSNT 212
QY 362 SMGTPIINDYTTKASGTSMATPHVSGVAGALLIOAHPSWTPDKVKTALITETADIAPKEIAD 421
Db 213 -----YTSINGTSMASPHVGAALILSKYPTLSAQVNRRLSTATNIGD-----S 259
QY 422 IAYGAGRVTYTKA 434
Db 260 FYYGKGLINVEAA 272

RESULT 10
US-10-541-737-3
/ Sequence 3, Application US/10541737
/ Publication No. US20060252155A1
/ GENERAL INFORMATION:
/ APPLICANT: Leeeflang, Chris
/ APPLICANT: Van der Kleij, Wilhelmus A.H.
/ TITLE OF INVENTION: Methods for Site-Directed Mutagenesis
/ FILE REFERENCE: GC787-2
/ CURRENT APPLICATION NUMBER: US/10/541,737
/ PRIOR FILING DATE: 2005-07-08
/ PRIOR APPLICATION NUMBER: PCT/US04/01334
/ PRIOR FILING DATE: 2004-01-16
/ PRIOR APPLICATION NUMBER: US 60/440,792
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: FaetsSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 274
/ TYPE: PRT
/ ORGANISM: Bacillus licheniformis
US-10-541-737-3

Query Match 15.2%; Score 521.5; DB 6; Length 274;
Best Local Similarity 42.2%; Pred. No. 1.5e-29;
Matches 132; Conservative 39; Mismatches 93; Indels 49; Gaps 11;

QY 124 GIKPIODYKVOYVDATSVSOGIADVTWNSLGYDGSVVAIVDTGIDANHPDLKGVIG 183
Db 7 GIPILIKAD-KVQAO-----GFKGANVKTAVLDGTGIAQHPDL--NVVG 46
QY 184 WYDAVNGRSTPYDQGHGTIVAGIAGTGSVNSQYIGVAPGAKLVGVKVLGADGSGSVT 243
Db 47 GASPVAGEAVNTDNGHGTIVAGTVALDNTTG-VLGAVPSVSLYAVKVLNSGSGSVSG 105
QY 244 IIAQVDMVQNKXGIRVNLISGSSQSDGTDSLQAVNNMADAIYVCVAAGNSG-- 301
Db 106 IYSGIEAATTN---GMDVINMSLG---GASGTAMQAVDNAARGVVVAAGNSGNS 158
QY 302 PNTYVTSPPAAASKVITVGVAVDSNDNIASFSSRGPTADGRLKPEVVAAPGVDIAPRASGT 361
Db 159 GSTTITIGPKYDSVIAVGAVDNSKNRASFSSVGS-----AELEVAAPG-----AGV 204
QY 362 SMGTPIINDYTTKASGTSMATPHVSGVAGALLIOAHPSWTPDKVKTALITETADIAPKEIAD 421
Db 205 YSTYPTVTYAT-INGTSMASPHVGAALILSKHPNLASQVNRRLSTATYIG-----SS 259
QY 422 IAYGAGRVTYTKA 434
Db 260 FYYGKGLINVEAA 272

RESULT 11
US-11-452-695-2
/ Sequence 2, Application US/11452695
/ Publication No. US20060222641A1
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Neil H. Riordan
/ TITLE OF INVENTION: METHOD AND COMPOSITION FOR PREVENTING OR
/ TITLE OF INVENTION: REDUCING DEEP VEIN THROMBOSIS AND/OR PULMONARY EMBOLISM
/ FILE REFERENCE: AIDAN.005DV1
/ CURRENT APPLICATION NUMBER: US/11/452,695
/ CURRENT FILING DATE: 2006-06-14
/ PRIOR APPLICATION NUMBER: 60/468948
/ PRIOR FILING DATE: 2003-05-07
/ PRIOR APPLICATION NUMBER: 10/647131
/ PRIOR FILING DATE: 2003-08-22
/ NUMBER OF SEQ ID NOS: 7
/ SOFTWARE: FaetsSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 381
/ TYPE: PRT
/ ORGANISM: Bacillus Subtilis
US-11-452-695-2

Query Match 15.0%; Score 516.5; DB 7; Length 381;
Best Local Similarity 38.2%; Pred. No. 5.3e-29;
Matches 139; Conservative 53; Mismatches 123; Indels 49; Gaps 13;

QY 78 YLRMGAQYKYSKILIPAVAVKIKARDLLIAGMIDTGYFGNTRVSGIKPIODYKVOYD 137
Db 58 VISEKQKVOQKQRYVNAATAATDEXAVKEL-----KQDPSVAYVEDH-IAHE 105
QY 138 DATSV-----SOGIADVTWNSLGYDGSVVAIVDTGIDANHPDLKGVIGWYDAVNGRST 193
Db 106 YASQVYVYGSQIAAPRL-HSQGTGSNVAVAVDSIDSHPL--NVAGGASFVSETN 162
QY 194 PYDD-QAGHTVAGIAGTGSVNSQYIGVAPGAKLVGVKVLGADGSGSVSTIAGYDWTY 252
Db 163 PYDQSSHGHTVAGTIAALNN-SIGVIGVAPASLAVKVLDTSTGSGQYSWIINGIEMAL 221
QY 253 QNKDKXGIRVYNLSGSSQSDGTDSLQAVNNMADAIYVCVAAGNSGP--NNTYVTS 310
Db 222 SN---NMDVINMSLG---GPTGSTALKTVYDAVSSGIVVAAAGNSGSGSTIVGTP 274
QY 311 AASKVITVGAVDNSNDNIASFSSRGPTADGRLKPEVVAAPGVDIAPRASGTSMGTPIINDY 370
Db 275 AKPSTIIVAGAVSSNQRASFSSVSGELD-----VMAPGVSIQSTLPQGT----- 319
QY 371 YTRASGTSMATPHVSGVAGALLIOAHPSWTPDKVKTALITETADIAPKEIADIAVYAGRVN 430
Db 320 YGAVNGTSMATPHVGAALILSKHPTWTAQVRDLSTATYIG-----SSFYTGGLIN 375
QY 431 VYKA 434
Db 376 VQAA 379

RESULT 12
US-11-452-695-3
/ Sequence 3, Application US/11452695
/ Publication No. US20060222641A1
/ GENERAL INFORMATION:
/ APPLICANT: Neil H. Riordan
/ TITLE OF INVENTION: METHOD AND COMPOSITION FOR PREVENTING OR
/ TITLE OF INVENTION: REDUCING DEEP VEIN THROMBOSIS AND/OR PULMONARY EMBOLISM
/ FILE REFERENCE: AIDAN.005DV1
/ CURRENT APPLICATION NUMBER: US/11/452,695
/ CURRENT FILING DATE: 2006-06-14
/ PRIOR APPLICATION NUMBER: 60/468948
/ PRIOR FILING DATE: 2003-05-07
/ PRIOR APPLICATION NUMBER: 10/647131
/ PRIOR FILING DATE: 2003-08-22
/ NUMBER OF SEQ ID NOS: 7
/ SOFTWARE: FaetsSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 381
/ TYPE: PRT
/ ORGANISM: Bacillus Subtilis
```

US-11-452-695-3

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 14.9%; | Score 513.5; | DB 7; | Length 381; |
| Best Local Similarity | 38.2%; | Pred. NO. 8.7e-29; | | |
| Matches 139; | Conservative 52; | Mismatches 124; | Indels 49; | Gaps 13; |

```

OY      78 VLRLMGAQVVKSYKILIPAVAVKIKARDDLILAGMIDTGYFNGTRVSGIKFIQEDYKVOVD 137
Db      58 VISKGGKQVQKQFCYVAAAATAATLDEKAVKEL-----KQDPVAIYBEDH-IAHE 105
OY      138 DATSV---SQIGADTVWNSLIGDYGSGVVVAIVDTGIDANHPDLKGVIGWYDANVGKST 193
Db      106 YAGSVPGYISQIKAPAL-HSQGYGTSNKVAVIBGIDSSHPLD--NVRGASFPVSETN 162
OY      194 PYDD-QGHGTHVAGIVACTGYSVNSOYIGVAAQAKLVKTVLGADSGSVSTIIAGVWV 252
Db      163 PYDGGSSHGHVACTIALANN-SIGVLGVAPASASIAKVLVDSTSGGYSWIINGIEVAI 221
OY      253 QNKDKYGRIVNLSLIGSSQSSDGTDSLISQAVNNMADAGIIVCVAAANGSGP--NTYVQSP 310
Db      222 SN---NMVDYINMSLG--GPTGSTALLTVYDQKAVSGIIVAAAANGSGSGSTIVGYR 274
OY      311 AAASKVITVGAVDSDNDNIASFSSRGPTADGRLKEBVAVPQVDIAPRASGTSMTGRPINDY 370
Db      275 AKYPTIATGAVNNSNQRASFSVSGSELD-----VNAFGVISQSTLPGT----- 319
OY      371 YTKASGTSMATPVHSGVGAALLLOHPSTPDKVKTALILETDIYAPKEIADIANGARVN 430
Db      320 YGANGTSMATPHVAGAAALILISKHPWTNAQVRDLESTATYLG---NSFYGKGGLIN 375
OY      431 YVKA 434
Db      376 VQAA 379

RESULT 13
US-11-452-695-4
; Sequence 4, Application US/11452695
; Publication No. US20060222641A1
; GENERAL INFORMATION:
; APPLICANT: Neil H. Riordan
; TITLE OF INVENTION: METHOD AND COMPOSITION FOR PREVENTING OR
; TITLE OF INVENTION: REDUCING DEEP VEIN THROMBOSIS AND/OR PULMONARY EMBOLISM
; FILE REFERENCE: AIDAN_005DVI
; CURRENT APPLICATION NUMBER: US/11/452,695
; CURRENT FILING DATE: 2006-06-14
; PRIOR APPLICATION NUMBER: 60/468948
; PRIOR FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: 10/647131
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Bacillus Subtilis
; US-11-452-695-4

```

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 14.9%; | Score 513.5; | DB 7; | Length 381; |
| Best Local Similarity | 38.2%; | Pred. No. 8.7e-29; | | |
| Matches 139; | Conservative 52; | Mismatches 124; | Indels 49; | Gaps 13; |

QY 78 VLRLMGAGQVRSYKIPVAVAKIVARDDLILAGMIDIGYFONTRVSGIKKIFIOEDVKQVDD 137

Db 58 VISEKGGKVKQCFKVVAAAAATLLEKAVKL-----KQPSVAAYEEHD-IHAE 105

QY 138 DATSV---SOIGADTVNWSLGYDGGSGVVAIVDPGIDANHPDLKAKIYGVDAWANGRST 193

Db 106 YAGSVPPVIGSIGIKPAL-HSGGYTGSNKKVAVIIISGIDSSHPLD--NVRGASFPVSEETN 162

QY 194 PYDD-QGHGHTVAGIVAGTGSVNSQYIGVAPGAKLVKRYKVDGADSGSVSTIIAGVDMVY 252

Db 163 PYQDGGSHGHTVACTIALNN-SIGVLGVAAPASALYAKKVLDTGSGGYSWINGIEVAI 221

| | | | |
|----|-----|--|--|
| QY | 255 | QNKDKKGGJFVNLSTLSSGSGSSDGDGDSLSQAVNNMADGDIYVCVAAAGSGP--NNTYTWGSP | 31.0 |
| | | 222 | SN---NMDDVIMSLG---GPTGSLALTYVDKAVSSGIVVAAAAAGEGSGSGSTIVGYP |
| QY | 311 | AAASKVITGAVDSNDNIIASFSSRGPTADGRLKEBVAAPGVDIIPAPRASGTMGTPINDY | 37.0 |
| Db | 275 | AKPSTPIANGAVNNSNQRAFSFSSVSGELD-----VMAPGVISQSTLPGGI----- | 31.9 |
| QY | 371 | YTKASGTSMAATHVSGVCGALLIQAHPSPTPKVKATALLETADYAPKEIADIANGARVN | 43.8 |
| Db | 320 | YGAVNGTSWATHVAGAAALILSKPIPTWNAQVEDRLESTATYLG---NSFYKGGLIN | 37.5 |
| QY | 431 | YVKA | 43.4 |
| Db | 376 | VOAA | 37.9 |

```

RESULT 14
US-10-526-324-20
? Sequence 20. Application US/10526324
? Publication NO. US20060248617A1
? GENERAL INFORMATION:
? APPLICANT: Imanaka, Takayuki
? APPLICANT: Atom, Haruyuki
? TITLE OF INVENTION: METHOD OF TARGETED GENE DISRUPTION, GENOME OF
? TITLE OF INVENTION: HYPERTHERMOSTABLE BACTERIUM AND GENOME CHIP USING
? TITLE OF INVENTION: THE SAME
? FILE REFERENCE: 490051.401USPC
? CURRENT FILING DATE: 2005-02-28
? PRIOR APPLICATION NUMBER: PCT/IB2003/003597
? PRIOR FILING DATE: 2003-08-29
? PRIOR APPLICATION NUMBER: JP 2002-319011
? PRIOR FILING DATE: 2002-08-30
? NUMBER OF SEQ ID NOS: 2167
? SOFTWARE: Patentin version 3.1
? SEQ ID NO 20
? LENGTH: 524
? TYPE: prt
? ORGANISM: Thermococcus kodakarensis KOD1
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (414542)..(414542)
? OTHER INFORMATION: n is a or c or g or t.
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (786890)..(786890)
? OTHER INFORMATION: n is a or c or g or t.
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (786907)..(786907)
? OTHER INFORMATION: n is a or c or g or t.
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (786944)..(786946)
? OTHER INFORMATION: n is a or c or g or t.
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (839139)..(839139)
? OTHER INFORMATION: n is a or c or g or t.
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1128488)..(1128488)
? OTHER INFORMATION: n is a or c or g or t.
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1128499)..(1128499)
? OTHER INFORMATION: n is a or c or g or t.
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1128505)..(1128506)
? OTHER INFORMATION: n is a or c or g or t.

```



```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1128517)..(1128518)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1128539)..(1128540)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1561400)..(1561400)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1561477)..(1561477)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1561545)..(1561545)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1767941)..(1767941)
; OTHER INFORMATION: n is a or c or g or t.
; US-10-526-324-20

```

Query Match 14.8%; Score 508.5; DB 6; Length 524;

Best Local Similarity 35.4%; Pred. No. 3.1e-28; Matches 146; Conservative 65; Mismatches 150; Indels 51; Gaps 12;

```

QY 73 DRAVKVRLMGAQVYKSIIPAVAVKIKARDLLIAGMIDTGFNGTRVSGIKFIOEDY 132
DB 124 DRKEALALGAELVYDDDDINAIKARPS---LKNLVYSAQFDFYRFTREVPDL 180
QY 133 KV-----QYDPAF-SYSGAGADTVNNSGYDSSGVVAIVDT 168
DB 181 YVETGPTILENETTTPPANTTTPPEQVSGAIWNKIKVKAOLAWSKEGITKGVVAIVDT 240
QY 169 GIDANHPDLKGVYIGWTDVANGRSTPYDDQGHGTHVAGVAG-----TGSVNSQYI-GVA 222
DB 241 GUDCDHYMLQACVGFNFPTDEPAK-DLNGHGHVAGITAGPRTKYTWGSKYVGVGA 239
QY 223 PGAKLVGVKVLGADGSGSVSTIIAGVDVYVONKDKYGR--VINLSIGSSQSSDCTDSL 280
DB 300 PEANIIIAVKVLGGDGGTGTQIQGLDYVEMKKHGBRPVISMISLSPGSPRDMVQ 359
QY 281 QAVNNMADAGIVCVANGSGPNTYVGSPPAASKVITTVGAVDNDNIIASFSSRGTAD- 339
DB 360 KVRQIIRREHIPPVIAAGN--EPAVIDSPGIAITVAADVNMKVASFSGKPGINT 416
QY 340 GRLKPEVAVAGVDIAPRASGTSWGTPIINDYTTKASGTSMATPHVSGVALLIQAHPSWT 399
DB 417 YDIKPIAAGVKILSKA-----GT--RHEFIAMSGTSMATPHVSGVALLIQAHPDWT 469
QY 400 PDKVKTALITADIAPKEIADIAVAGAVNVYKAIKYDDYAKLTFTGSVAD 451
DB 470 PETIKLILKTAIYPLDGDIALPTWSGAGVVDAYAAVK---AEPSESGLID 517

```

```

RESULT 15
US-11-433-614-16
; Sequence 16, Application US/11433614
; Publication No. US20060205019A1
; GENERAL INFORMATION:
; APPLICANT: Athena Biotechnologies, Inc.
; APPLICANT: Menzel, Rolf
; TITLE OF INVENTION: Methods and Compositions for Directed Gene Assembly
; FILE REFERENCE: 46675-5004-01-US
; CURRENT APPLICATION NUMBER: US/11/433.614
; PRIORITY FILING DATE: 2006-05-12
; PRIORITY APPLICATION NUMBER: US 09/920,118
; PRIORITY FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 22

```

```

; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16
; LENGTH: 381
; TYPE: PRF
; ORGANISM: Bacillus subtilis
US-11-433-614-16

```

Query Match 14.8%; Score 507.5; DB 7; Length 381;

Best Local Similarity 37.9%; Pred. No. 2.3e-28; Matches 138; Conservative 52; Mismatches 125; Indels 49; Gaps 13;

```

QY 78 VLRLMGAQVYKSIIPAVAVKIKARDLLIAGMIDTGFNGTRVSGIKFIOEDYKQVD 137
DB 58 VISEKGVQKQKPKVYNAATAATDEKAVEL-----KKDSVAIVEEDH-IAHE 105
QY 138 DATSV-----SQIGADTVNNSLGYDSSGVVAIVDTGIDANHPDLKGVIGWTDVANGRST 193
DB 106 YAGSVYIGISQIKAPAL-HSQGTGSNVKVAIVDSIGDSSHPL--NVAGASFFVSEFN 162
QY 194 PYDD-QGHGTHVAGIVAGTGSVNSQYIVAPGAKLVGVKVLGADGSGSVSTIIAGVDVYV 252
DB 163 PYDDGSSHGTHVAGTIAALNN-SIGVLGVAAPSALYAVKVLDTSTGSGQYSWIINGIEMAI 221
QY 253 QNKDKYGIKVINLSIGSSQSSDCTDSLGAANNMADAGIVCVANGSGP--NTYVGSF 310
DB 222 SN---NMVYNNISL--GPTGSTALKTVVDRAVSSGIVAAAAGNBSGSGSTIVGIF 274
QY 311 AAASKVITVGAVDSDNIIASFSSRGTADGRLKPEVAVAGVDIAPRASGTSWGTPIINDY 370
DB 275 AKYPSIIAVGAVVNSQNRASFSSAGSELD-----VMAQGVSIQSTLPQGT----- 319
QY 371 YTKASGTSMATPHVSGVALLIQAHPSWTBDKYKTLITADIAPKEIADIAVAGAVN 430
DB 320 YGAVNGTSMATPHVARAALILISKHPTWNAQVRDLSEATYLG-----NSFYTGKGLIN 375
QY 431 VYKA 434
DB 376 VQAA 379

```

Search completed: January 6, 2007, 22:39:01

Job time : 33 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2007 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2007, 22:19:02 ; Search time 142 Seconds
(without alignments)
2121.871 Million cell updates/sec

Title: US-10-800-684-5

Perfect score: 3428
Sequence: 1 MKGLKALIVLVGLVVGVS.....YASTYGMADYQLKAVVYVG 659

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.8:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20038:*
8: geneseqp20048:*
9: geneseqp20058:*
10: geneseqp20068:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|---------------------|
| 1 | 3428 | 100.0 | 659 | 2 | AAW24123 Protease. |
| 2 | 2997.5 | 87.4 | 654 | 2 | AAW24129 Pyrococcus |
| 3 | 2997.5 | 87.4 | 654 | 2 | AAW94841 Hyperther |
| 4 | 2914 | 85.0 | 659 | 2 | AAW24121 Thermococ |
| 5 | 2914 | 85.0 | 659 | 2 | AAW94840 W09856926 |
| 6 | 2705 | 78.9 | 663 | 8 | ADNA6373 Thermococ |
| 7 | 2349.5 | 68.5 | 522 | 2 | AAW24122 Pyrococcus |
| 8 | 2349.5 | 68.5 | 522 | 2 | AAW94838 Hyperther |
| 9 | 1918 | 56.0 | 412 | 2 | AAW94836 Hyperther |
| 10 | 1051.5 | 30.7 | 237 | 2 | AAW87009 Hyperther |
| 11 | 909 | 26.5 | 545 | 4 | ABW09483 T. yonsei |
| 12 | 864 | 25.2 | 168 | 2 | AAW24127 Pyrococcus |
| 13 | 862.5 | 25.2 | 188 | 2 | AAW87011 Peptide S |
| 14 | 862.5 | 25.2 | 188 | 2 | AAW24126 Thermococ |
| 15 | 692.5 | 20.2 | 734 | 2 | AAW13667 Streptomy |
| 16 | 692.5 | 20.2 | 823 | 2 | AAW13668 DHPA-mel |
| 17 | 679.5 | 19.8 | 520 | 2 | AAW13666 Fragment |
| 18 | 670 | 19.5 | 443 | 8 | ADM99125 Bacterial |
| 19 | 662.5 | 19.3 | 1079 | 4 | AAW81180 Transglut |
| 20 | 660 | 19.3 | 1079 | 6 | ABU07391 Foreign p |
| 21 | 655 | 19.1 | 1237 | 6 | ADM99123 Bacterial |
| 22 | 655 | 19.1 | 1237 | 6 | ABU11343 Protein e |
| 23 | 588 | 17.2 | 552 | 8 | ADM99241 Envelope |

| | | | | | |
|----|-------|------|-----|----|---------------------|
| 24 | 586 | 17.1 | 182 | 4 | ABW09484 T. yonsei |
| 25 | 565 | 16.5 | 382 | 5 | ABW79087 Mutant su |
| 26 | 562 | 16.4 | 379 | 6 | AAW29944 Bacillus |
| 27 | 561 | 16.4 | 372 | 6 | AAW29942 Bacillus |
| 28 | 561 | 16.4 | 382 | 6 | ADA50345 B. amyloli |
| 29 | 561 | 16.4 | 383 | 5 | ADW54158 L-25 kera |
| 30 | 560 | 16.3 | 379 | 10 | AEF11989 Bacillus |
| 31 | 558 | 16.3 | 379 | 2 | AAW04137 Sequence |
| 32 | 558 | 16.3 | 382 | 5 | ABW79088 Mutant su |
| 33 | 557.5 | 16.3 | 373 | 2 | AAW86879 Subtilisi |
| 34 | 557 | 16.2 | 379 | 5 | AAW88521 Keratins |
| 35 | 557 | 16.2 | 379 | 5 | AAW19062 Bacillus |
| 36 | 556.5 | 16.2 | 394 | 2 | AAW34261 Bacillus |
| 37 | 556 | 16.2 | 379 | 2 | AAW26762 Bacillus |
| 38 | 555 | 16.2 | 382 | 1 | AAW50880 Sequence |
| 39 | 555 | 16.2 | 382 | 1 | AAW70486 Bacillus |
| 40 | 555 | 16.2 | 382 | 1 | AAW80271 Amino aci |
| 41 | 555 | 16.2 | 382 | 2 | AAW58691 Subtilisi |
| 42 | 555 | 16.2 | 382 | 2 | AAW75161 B. amylo |
| 43 | 555 | 16.2 | 382 | 2 | AAW65247 B. amylo |
| 44 | 555 | 16.2 | 382 | 2 | AAW74223 B. amylo |
| 45 | 555 | 16.2 | 382 | 2 | AAW86876 Subtilisi |

ALIGNMENTS

RESULT 1
ID AAW24123 standard; protein; 659 AA.
XX AAW24123;
AC AAW24123;
XX 20-APR-1998 (first entry)
DT 20-APR-1998 (first entry)
DE Protease.
XX Protease.
XX Protease; research reagent; thermal stability.
KM Protease; research reagent; thermal stability.
XX Synthetic.
XX WO9721823-A1.
XX 19-JUN-1997.
PD 19-JUN-1997.
XX 07-NOV-1996; 96WO-JP003253.
PF 07-NOV-1996; 96WO-JP003253.
XX 12-DEC-1995; 95UP-00323285.
PR 12-DEC-1995; 95UP-00323285.
XX (TAKI) TAKARA SHUZO CO LTD.
PA (TAKI) TAKARA SHUZO CO LTD.
XX Takakura H, Morishita M, Yamamoto K, Mitra M, Asada K;
PI Tsunashawa S, Kato I;
XX WPI; 1997-332794/30.
DR N-PSDB; AAT85669.
XX Protease (s) and genes encoding them obtained from Thermococcus and
PT Pyrococcus strains - have extremely high thermal stability and are useful
XX industrially and as research reagents.
XX Claim 9; Page 92-95; 159pp; Japanese.
PS This sequence is a protease of the invention. The proteases of the
XX invention have extremely high thermal stability. The proteases can be
CC used as research reagents, and industrially in the food, drug and
CC chemical industries
XX Sequence 659 AA;
SQ
Query Match 100.0%; Score 3428; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 1.2e-208;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MKGLKALLIIVILVGLVGSVAAAPKKVEQVRNVEKNYGLTPGLFRKIQKLNPEEIS 60
DB 1 MKGLKALLIIVILVGLVGSVAAAPKKVEQVRNVEKNYGLTPGLFRKIQKLNPEEIS 60
QY 61 TVIVFENHREKEIAVRVLELMGAKRVYVHHIIPALAADLKVRDLVLSGLTGKAKLSGV 120
DB 61 TVIVFENHREKEIAVRVLELMGAKRVYVHHIIPALAADLKVRDLVLSGLTGKAKLSGV 120
QY 121 RFIQEDKYTVSAELGDESAQAQVATYVMNLGYDGSGITIGIDTGDASHPDLQSKV 180
DB 121 RFIQEDKYTVSAELGDESAQAQVATYVMNLGYDGSGITIGIDTGDASHPDLQSKV 180
QY 181 IGMWDFVNGRSYFYDDHGHGTHVASTAAGTGAASNGKRYGMAPGAKLAGIKVLGADSGS 240
DB 181 IGMWDFVNGRSYFYDDHGHGTHVASTAAGTGAASNGKRYGMAPGAKLAGIKVLGADSGS 240
QY 241 ISTTIKGVMAVDNKKYGIKVINLSIGSSQSSDGTDSLQAVNNAMDAGIIVCVAAAGNS 300
DB 241 ISTTIKGVMAVDNKKYGIKVINLSIGSSQSSDGTDSLQAVNNAMDAGIIVCVAAAGNS 300
QY 301 GPMYTVVGSPPAASKVITVGAVDSDNINASFSSRGPTADGRLEPEVAPGVDIIPARASG 360
DB 301 GPMYTVVGSPPAASKVITVGAVDSDNINASFSSRGPTADGRLEPEVAPGVDIIPARASG 360
QY 361 TSMGTPINDYTTKASGTSMAATPHVSGVGLILQAHPSWTPDKYKTALIEIADIVAPKEIA 420
DB 361 TSMGTPINDYTTKASGTSMAATPHVSGVGLILQAHPSWTPDKYKTALIEIADIVAPKEIA 420
QY 421 DIAVAGRVNVYKAIKYDDYAKLTFGTSVADKGSATHTFVSGATFTATLTYMDTGSDDI 480
DB 421 DIAVAGRVNVYKAIKYDDYAKLTFGTSVADKGSATHTFVSGATFTATLTYMDTGSDDI 480
QY 481 DLFLYDPNGNEVDYSYATAYGFEKVGYNPTAGTWTVKVSYKGAANYQVDVSDGSLQ 540
DB 481 DLFLYDPNGNEVDYSYATAYGFEKVGYNPTAGTWTVKVSYKGAANYQVDVSDGSLQ 540
QY 541 SGGGNPNPNPNPTPTTDTOTFTGSVNDYMDSDTFMNMNSGATKITGDLTFDTSYND 600
DB 541 SGGGNPNPNPNPTPTTDTOTFTGSVNDYMDSDTFMNMNSGATKITGDLTFDTSYND 600
QY 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPAPGTMFLVYASTYGADYQLKAVVYGG 659
DB 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPAPGTMFLVYASTYGADYQLKAVVYGG 659

```

RESULT 2

AAW24129 standard; protein; 654 AA.

```

XX AAW24129;
AC 17-OCT-2003 (revised)
DT 20-APR-1998 (first entry)
XX Pyrococcus furiosus protease PFUS.
KM Protease; research reagent; thermal stability; pyrococcus furiosus.
XX Pyrococcus furiosus; DSM-3638.
OS MO9721823-A1.
XX 19-JUN-1997.
XX 07-NOV-1996; 96MO-JP003253.
XX 12-DEC-1995; 95JP-00323285.
XX (TAKI ) TAKARA SHUZO CO LTD.
XX Takakura H, Morishita M, Yamamoto K, Mita M, Asada K;
PI Tsunashawa S, Kato I;

```

```

XX WPI: 1997-332794/30.
DR N-PSDB; AAT85635.
XX Protease(s) and genes encoding them obtained from Thermococcus and
PT Pyrococcus strains - have extremely high thermal stability and are useful
PT industrially and as research reagents.
XX Disclosure; Page 125-128; 159pp; Japanese.
XX This sequence represents the protease from Pyrococcus furiosus DSM-3638.
CC This sequence is a protease of the invention. The proteases of the
CC invention have extremely high thermal stability. The proteases can be
CC used as research reagents, and industrially in the food, drug and
CC chemical industries. (updated on 17-OCT-2003 to standardise Os field)
XX
SQ Sequence 654 AA;

```

Query Match 87.4%; Score 2997.5; DB 2; Length 654;
 Best Local Similarity 88.3%; Pred. No. 2.3e-181;
 Matches 582; Conservative 22; Mismatches 50; Indels 5; Gaps 1;

```

QY 1 MKGLKALLIIVILVGLVGSVAAAPKKVEQVRNVEKNYGLTPGLFRKIQKLNPEEIS 60
DB 1 MKGLKALLIIVILVGLVGSVAAAPKKVEQVRNVEKNYGLTPGLFRKIQKLNPEEIS 60
QY 61 TVIVFENHREKEIAVRVLELMGAKRVYVHHIIPALAADLKVRDLVLSGLTGKAKLSGV 120
DB 61 TVIVFENHREKEIAVRVLELMGAKRVYVHHIIPALAADLKVRDLVLSGLTGKAKLSGV 120
QY 121 RFIQEDKYTVSAELGDESAQAQVATYVMNLGYDGSGITIGIDTGDASHPDLQSKV 180
DB 121 RFIQEDKYTVSAELGDESAQAQVATYVMNLGYDGSGITIGIDTGDASHPDLQSKV 180
QY 181 IGMWDFVNGRSYFYDDHGHGTHVASTAAGTGAASNGKRYGMAPGAKLAGIKVLGADSGS 240
DB 181 IGMWDFVNGRSYFYDDHGHGTHVASTAAGTGAASNGKRYGMAPGAKLAGIKVLGADSGS 240
QY 241 ISTTIKGVMAVDNKKYGIKVINLSIGSSQSSDGTDSLQAVNNAMDAGIIVCVAAAGNS 300
DB 241 ISTTIKGVMAVDNKKYGIKVINLSIGSSQSSDGTDSLQAVNNAMDAGIIVCVAAAGNS 300
QY 301 GPMYTVVGSPPAASKVITVGAVDSDNINASFSSRGPTADGRLEPEVAPGVDIIPARASG 360
DB 301 GPMYTVVGSPPAASKVITVGAVDSDNINASFSSRGPTADGRLEPEVAPGVDIIPARASG 360
QY 361 TSMGTPINDYTTKASGTSMAATPHVSGVGLILQAHPSWTPDKYKTALIEIADIVAPKEIA 420
DB 361 TSMGTPINDYTTKASGTSMAATPHVSGVGLILQAHPSWTPDKYKTALIEIADIVAPKEIA 420
QY 421 DIAVAGRVNVYKAIKYDDYAKLTFGTSVADKGSATHTFVSGATFTATLTYMDTGSDDI 480
DB 421 DIAVAGRVNVYKAIKYDDYAKLTFGTSVADKGSATHTFVSGATFTATLTYMDTGSDDI 480
QY 481 DLFLYDPNGNEVDYSYATAYGFEKVGYNPTAGTWTVKVSYKGAANYQVDVSDGSLQ 540
DB 481 DLFLYDPNGNEVDYSYATAYGFEKVGYNPTAGTWTVKVSYKGAANYQVDVSDGSLQ 540
QY 541 SGGGNPNPNPNPTPTTDTOTFTGSVNDYMDSDTFMNMNSGATKITGDLTFDTSYND 600
DB 541 SGGGNPNPNPNPTPTTDTOTFTGSVNDYMDSDTFMNMNSGATKITGDLTFDTSYND 600
QY 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPAPGTMFLVYASTYGADYQLKAVVYGG 659
DB 596 LDLYLDPNGNLVDRSTSSNSYEHVEYANPAPGTMFLVYASTYGADYQLKAVVYGG 654

```

RESULT 3

AAW94841 standard; protein; 654 AA.

```

XX AAW94841;
AC
XX

```

DT 06-MAY-1999 (first entry)
 XX Hyperthermostable protease.
 DE Hyperthermostable protease.
 XX
 KM Hyperthermostable; protease; thermophilic; bacterium; subtilisin;
 KM additive; drug; washing agent; foodstuff; chemical synthesis.
 XX
 OS *Pyrococcus furiosus*.
 XX
 PN M09856926-A1.
 XX
 PD 17-DEC-1998.
 XX
 PF 04-JUN-1998; 98WO-JP002465.
 XX
 PR 10-JUN-1997; 97JP-00151969.
 XX
 PA (TAKI) TAKARA SHUZO CO LTD.
 XX
 PI Takakura H, Morishita M, Shimojo T, Asada K, Kato I;
 XX
 DR WPI: 1999-080907/07.
 DR N-PSDB; AAX05929.
 XX
 PT Recombinant hyperthermostable protease from *Pyrococcus furiosus* - and
 PT gene encoding it, for large scale production of the protease for
 PT industrial use.
 XX
 PS Disclosure; Page 60-63; 82pp; Japanese.
 XX
 CC The invention relates to a hyperthermostable protease derived from a
 CC thermophilic bacterium (especially *Pyrococcus furiosus*). The protease has
 CC working temperature 40-110 deg.C (optimum 80-95 deg.C), working pH 5-10
 CC (optimum 6-8), and retains more than 90% of its activity after 8 hours at
 CC 95 deg.C. The invention also provides gene sequences encoding a
 CC polypeptide of formula SIG-Ala-Gly-Asn-PRO, where SIG is a signal
 CC peptide from subtilisin, and PRO is the above protease. Host cells
 CC (especially *Bacillus* strains) transformed with vectors comprising the
 CC genes are used for the recombinant production of the protease. The
 CC hyperthermostable protease which can be prepared in quantity suitable for
 CC industrial use, can be used as an additive for drugs, washing agents and
 CC foodstuffs and for chemical synthesis
 XX
 SQ Sequence 654 AA;

Query Match 87.4%; Score 2997.5; DB 2; Length 654;
 Best Local Similarity 88.3%; Pred. No. 2.3e-181;
 Matches 582; Conservative 22; Mismatches 50; Indels 5; Gaps 1;

QY 1 MKGLKALLIVLVGLVGVSVAAAPKRYQVNRKRYGGLTRGLFRKIQKLNPNERS 60
 DB 1 MKGLKALLIVLVGLVGVSVAAAPKRYQVNRKRYGGLTRGLFRKIQKLNPNERS 60
 QY 61 TYIVFENHREKEIAVAVLELMGAKVYVYHIIPALAADLKVRDLVTSGLTGKATLGSV 120
 DB 61 TYIVFENHREKEIAVAVLELMGAKVYVYHIIPALAADLKVRDLVTSGLTGKATLGSV 120
 QY 61 TYIVFENHREKEIAVAVLELMGAKVYVYHIIPALAADLKVRDLVTSGLTGKATLGSV 120
 DB 61 TYIVFENHREKEIAVAVLELMGAKVYVYHIIPALAADLKVRDLVTSGLTGKATLGSV 120
 QY 121 RFIQEDYKVTSALEGLDESAAQVMATYVWNLGYDGSGITIGI IDTGIDASHPLDQSKY 180
 DB 121 RFIQEDYKVTSALEGLDESAAQVMATYVWNLGYDGSGITIGI IDTGIDASHPLDQSKY 180
 QY 121 RFIQEDYKVTSALEGLDESAAQVMATYVWNLGYDGSGITIGI IDTGIDASHPLDQSKY 180
 DB 121 RFIQEDYKVTSALEGLDESAAQVMATYVWNLGYDGSGITIGI IDTGIDASHPLDQSKY 180
 QY 181 IGMVDPVNRSPYRDHGHGTHVASTAAGTGAASNKRYGMAFGALAGIKVLGADGSSG 240
 DB 181 IGMVDPVNRSPYRDHGHGTHVASTAAGTGAASNKRYGMAFGALAGIKVLGADGSSG 240
 QY 241 ISTIIIGVMAVNDKQYKIGKVINLSLGSQSSDGTDSLQAVNNAMADAGIVVCVAAGNS 300
 DB 241 ISTIIIGVMAVNDKQYKIGKVINLSLGSQSSDGTDSLQAVNNAMADAGIVVCVAAGNS 300
 QY 301 GPNTTYVGSPPAAASKITTYGAVDSNDNIAFSRSGPTADGRLEPVAAPGVDTIAPRAGS 360
 DB 301 GPNTTYVGSPPAAASKITTYGAVDSNDNIAFSRSGPTADGRLEPVAAPGVDTIAPRAGS 360
 QY 301 GPNTTYVGSPPAAASKITTYGAVDSNDNIAFSRSGPTADGRLEPVAAPGVDTIAPRAGS 360
 DB 301 GPNTTYVGSPPAAASKITTYGAVDSNDNIAFSRSGPTADGRLEPVAAPGVDTIAPRAGS 360

QY 361 TSMGTPIINDYTTKASGTSMAFPHVSGVALLIQAHPSWTPDKYKTALITADIIVAPKEIA 420
 DB 361 TSMGTPIINDYTTKASGTSMAFPHVSGVALLIQAHPSWTPDKYKTALITADIIVAPKEIA 420
 QY 421 DIAYGAGRNVYKAIKIDYALTLFTGSAVDGSAHTHTDVAGATVNTLTWDTGSSDI 480
 DB 421 DIAYGAGRNVYKAIKIDYALTLFTGSAVDGSAHTHTDVAGATVNTLTWDTGSSDI 480
 QY 481 DLYLYDPNGNEVDYATAYGFEKVGYNPTAGTWTVKVSYKGAANYOVVYVSDGSLQ 540
 DB 481 DLYLYDPNGNEVDYATAYGFEKVGYNPTAGTWTVKVSYKGAANYOVVYVSDGSLQ 540
 QY 541 SGGGNPNPNPNPNPTPTDTGTFTGSVNDYMDTSDTFMNNVNSGATKIGDLTFDTSYND 600
 DB 541 SGGGNPNPNPNPNPTPTDTGTFTGSVNDYMDTSDTFMNNVNSGATKIGDLTFDTSYND 600
 QY 601 LDLYLYDPNGNLVDNSTSNSTYEHVEYANPAQGTWFLVYAASTYGMADYQLKAVYYG 659
 DB 596 LDLYLYDPNGNLVDNSTSNSTYEHVEYANPAQGTWFLVYAASTYGMADYQLKAVYYG 659

RESULT 4
 AAM24121
 ID AAM24121 standard; protein; 659 AA.
 XX
 AC AAM24121;
 XX
 DT 17-OCT-2003 (revised)
 XX
 DT 20-APR-1998 (first entry)
 XX
 DE Thermococcus protease.
 XX
 KM Protease; research reagent; thermal stability; thermococcus celer.
 XX
 OS Thermococcus celer; DSM-2476.
 XX
 PN M09721823-A1.
 XX
 PD 19-JUN-1997.
 XX
 PF 07-NOV-1996; 96WO-JP003253.
 XX
 PR 12-DEC-1995; 95JP-00323285.
 XX
 PA (TAKI) TAKARA SHUZO CO LTD.
 XX
 PI Takakura H, Morishita M, Yamamoto K, Mita M, Asada K;
 XX
 DR WPI: 1997-332794/30.
 DR N-PSDB; AAT85667.
 XX
 PT Protease(s) and genes encoding them obtained from *Thermococcus* and
 PT *Pyrococcus* strains - have extremely high thermal stability and are useful
 PT industrially and as research reagents.
 XX
 PS Claim 1; Page 82-85; 159pp; Japanese.
 XX
 CC This sequence represents the protease from *Thermococcus celer* DSM-2476.
 CC This sequence is a protease of the invention. The proteases of the
 CC invention have extremely high thermal stability. The proteases can be
 CC used as research reagents, and industrially in the food, drug and
 CC chemical industries. (Updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 659 AA;

Query Match 85.0%; Score 2914; DB 2; Length 659;
 Best Local Similarity 83.7%; Pred. No. 4.5e-176;
 Matches 558; Conservative 47; Mismatches 46; Indels 16; Gaps 7;

QY 1 MKGLKALLIVLVGLVGVSVAAAPKRYQVNRKRYGGLTRGLFRKIQKLNPNERS 57
 DB 1 MKGLKALLIVLVGLVGVSVAAAPKRYQVNRKRYGGLTRGLFRKIQKLNPNERS 57

```

QY 58 EISTVIVENHREKEIAVRVLELMGAKRYVYHIIIPAAADLKVRDLVLSGL--TG--G 113
DB 59 EVDIVIMFGSGDSDRAVKVRLMGAQVKYSKIIPAAVAKIKARDLLIAGMDITGYFG 118
QY 114 KAKLSGVRFIOEDYKVTYSALEGLDESAQVMATYVWN-LGYDGSGITIGIIDTGIDAS 172
DB 119 NTRVSGIKFTIOEDYKVVDDA-----TSVSGIGADVTWNSISGYDGSVVVAIVDTGIDAN 173
QY 173 HPDLQKVIQWVDVFNRSYRYDDHGHGTVAASIAAGTGAASNGKYGMAGKLAGIKY 232
DB 174 HPDLKGVIGYDVAVNGRSTFYDQGHGTHVAGIVAGTGSV-NSQYIGVAPGAKLVGVKV 232
QY 233 LGADSGSISTIIKGVMAVDNKKDKYIKYINLSLSSGSSDGTDSLQAVNNAMDAGIV 232
DB 233 LGADSGSVSTIIAGVDMVQNKOKYGRIRVINSLSGSSSDGTDSLQAVNNAMDAGIV 232
QY 293 VCVAAAGNSGPTTYTGVSPAAASKVITYGAVDSDNNTASFSSRGPTADGRLEPVVAPGV 352
DB 293 VCVAAAGNSGPTTYTGVSPAAASKVITYGAVDSDNNTASFSSRGPTADGRLEPVVAPGV 352
QY 353 IIPRASGTMGPTINDYYTKASGTSMAATPHVSGVGLIIQAHPSWTPDKVKTALIEAD 412
DB 353 IIPRASGTMGPTINDYYTKASGTSMAATPHVSGVGLIIQAHPSWTPDKVKTALIEAD 412
QY 413 IVAPKEIADIVAGRVVYKAIKYDDYAKLFTGSVADKSGATHTFEDVSGATFVATLY 472
DB 413 IVAPKEIADIVAGRVVYKAIKYDDYAKLFTGSVADKSGATHTFEDVSGATFVATLY 472
QY 473 WDTGSSSIDLYLVDPNNGNEVDYSTAYYGFEEKYGYNPTAGTWTFLVYASTYGMADYOL 532
DB 473 WDTGSSSIDLYLVDPNNGNEVDYSTAYYGFEEKYGYNPTAGTWTFLVYASTYGMADYOL 532
QY 533 VSDGSLSSGSGGNPNPNPPTPTDQTFGTGSVNDYMDTSDPTFMNVNSGATKITGDL 592
DB 533 VSDGSLSSGSGGNPNPNPPTPTDQTFGTGSVNDYMDTSDPTFMNVNSGATKITGDL 592
QY 593 TFDTSYNDLIDLVLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLVYASTYGMADYOL 652
DB 593 TFDTSYNDLIDLVLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLVYASTYGMADYOL 652
QY 653 KAVVYVG 659
DB 653 KAVVYVG 659

```

RESULT 5
AAM94840
AAM94840 standard; protein; 659 AA.

AC AAM94840;
DT 06-MAY-1999 (first entry)
DE MO9856926 Seq ID 12.
XX MO9856926-AL.
XX 17-DEC-1998.
XX 04-JUN-1998; 98MO-JP002465.
XX 10-JUN-1997; 97JP-0015169.
XX (TAKI) TAKARA SHUZO CO LTD.
XX Takakura H, Morishita M, Shimojo T, Asada K, Kato I;

```

DR WPI: 1999-080907/07.
DR N-PSDB; AAX05926.
PT Recombinant hyperthermostable protease from Pyrococcus furiosus - and
PT gene encoding it, for large scale production of the protease for
PT industrial use.
PS Disclosure; Page 55-58; 82pp; Japanese.
XX
CC The invention relates to a hyperthermostable protease derived from a
CC thermophilic bacterium (especially Pyrococcus furiosus). The protease has
CC working temperature 40-110 deg.C (optimum 80-95 deg.C), working pH 5-10
CC (optimum 6-8), and retains more than 90% of its activity after 8 hours at
CC 95 deg.C. The invention also provides gene sequences encoding a
CC polypeptide of formula SIG-A1a-Gly-Gly-Aan-PRO, where SIG is a signal
CC peptide from subtilisin, and PRO is the above protease. Host cells
CC (especially Bacillus strains) transformed with vectors comprising the
CC genes are used for the recombinant production of the protease. The
CC hyperthermostable protease which can be prepared in quantity suitable for
CC industrial use, can be used as an additive for drugs, washing agents and
CC foodstuffs and for chemical synthesis
SQ
Sequence 659 AA;
Query Match 85.0%; Score 2914; DB 2; Length 659;
Best local Similarity 83.7%; Pred. No. 4, 5e-176;
Matches 558; Conservative 47; Mismatches 46; Indels 16; Gaps 7;
QY 1 MKGLALILVILVLGLVVGSAVAAPKRYBOVRN--VERKYGILLTTPGLFRKIQTLPNPE 57
DB 1 MKRLGAVVLAIVLVGLLGLTALAAVPRV--VRNNAVQOKYKGLLTPLGFKVQGMNMNQ 58
QY 58 EISTVIVENHREKEIAVRVLELMGAKRYVYHIIIPAAADLKVRDLVLSGL--TG--G 113
DB 59 EVDIVIMFGSGDSDRAVKVRLMGAQVKYSKIIPAAVAKIKARDLLIAGMDITGYFG 118
QY 114 KAKLSGVRFIOEDYKVTYSALEGLDESAQVMATYVWN-LGYDGSGITIGIIDTGIDAS 172
DB 119 NTRVSGIKFTIOEDYKVVDDA-----TSVSGIGADVTWNSISGYDGSVVVAIVDTGIDAN 173
QY 173 HPDLQKVIQWVDVFNRSYRYDDHGHGTVAASIAAGTGAASNGKYGMAGKLAGIKY 232
DB 174 HPDLKGVIGYDVAVNGRSTFYDQGHGTHVAGIVAGTGSV-NSQYIGVAPGAKLVGVKV 232
QY 233 LGADSGSISTIIKGVMAVDNKKDKYIKYINLSLSSGSSDGTDSLQAVNNAMDAGIV 292
DB 233 LGADSGSVSTIIAGVDMVQNKOKYGRIRVINSLSGSSSDGTDSLQAVNNAMDAGIV 292
QY 293 VCVAAAGNSGPTTYTGVSPAAASKVITYGAVDSDNNTASFSSRGPTADGRLEPVVAPGV 352
DB 293 VCVAAAGNSGPTTYTGVSPAAASKVITYGAVDSDNNTASFSSRGPTADGRLEPVVAPGV 352
QY 353 IIPRASGTMGPTINDYYTKASGTSMAATPHVSGVGLIIQAHPSWTPDKVKTALIEAD 412
DB 353 IIPRASGTMGPTINDYYTKASGTSMAATPHVSGVGLIIQAHPSWTPDKVKTALIEAD 412
QY 413 IVAPKEIADIVAGRVVYKAIKYDDYAKLFTGSVADKSGATHTFEDVSGATFVATLY 472
DB 413 IVAPKEIADIVAGRVVYKAIKYDDYAKLFTGSVADKSGATHTFEDVSGATFVATLY 472
QY 473 WDTGSSSIDLYLVDPNNGNEVDYSTAYYGFEEKYGYNPTAGTWTFLVYASTYGMADYOL 532
DB 473 WDTGSSSIDLYLVDPNNGNEVDYSTAYYGFEEKYGYNPTAGTWTFLVYASTYGMADYOL 532
QY 533 VSDGSLSSGSGGNPNPNPPTPTDQTFGTGSVNDYMDTSDPTFMNVNSGATKITGDL 592
DB 533 VSDGSLSSGSGGNPNPNPPTPTDQTFGTGSVNDYMDTSDPTFMNVNSGATKITGDL 592
QY 593 TFDTSYNDLIDLVLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLVYASTYGMADYOL 652
DB 593 TFDTSYNDLIDLVLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLVYASTYGMADYOL 652
QY 653 KAVVYVG 659

```


PI Takakura H, Morishita M, Yamamoto K, Mitra M, Asada K,
PI Tsumasawa S, Kato I;
DR WPI, 1997-332794/30.
XX N-PSDB; AAT85668.

PT Pyrococcus (s) and genes encoding them obtained from Thermococcus and
PT Pyrococcus strains - have extremely high thermal stability and are useful
XX industrially and as research reagents.

PS Claim 5; Page 87-90; 159pp; Japanese.

XX This sequence represents the protease from Pyrococcus furiosus DSM-3638.
CC This sequence is a protease of the invention. The proteases of the
CC invention have extremely high thermal stability. The proteases can be
CC used as research reagents, and industrially in the food, drug and
CC chemical industries. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 522 AA;

Query Match 68.5%; Score 2349.5; DB 2; Length 522;
Best Local Similarity 85.2%; Pred. No. 2e-140;
Matches 449; Conservative 22; Mismatches 51; Indels 5; Gaps 1;

QY 133 AELBGLDESSAAQWATYVMNLGYDGSGITIGITIDGIDASHPDLOGKVTGWDFNGRSY 192
DB 1 AELBGLDESSAAQWATYVMNLGYDGSGITIGITIDGIDASHPDLOGKVTGWDFNGRSY 60
QY 193 PYDDHGHGHTHVASIAAGTGAASNGKYKGMAPGAKIAGIKVLGADGSGSISTIIKGYEMAV 252
DB 61 PYDDHGHGHTHVASIAAGTGAASNGKYKGMAPGAKIAGIKVLGADGSGSISTIIKGYEMAV 120
QY 253 DNKDKYGIKVINLSLSSGSSDGTDSLISQAVNNAMDAGIYVVAAGNSGPNYTVGSPPA 312
DB 121 DNKDKYGIKVINLSLSSGSSDGTDSLISQAVNNAMDAGIYVVAAGNSGPNYTVGSPPA 180
QY 313 ASKYITVGAVDNSNNIASFSSRGPTADGRKPEVVAAGVUIIAPRASGTSMTGPIINDYTT 372
DB 181 ASKYITVGAVDKDVITFSFSSRGPTADGRKPEVVAAGVUIIAPRASGTSMTGPIINDYTT 240
QY 373 KASGTSMATPHVSGVGLILQAHPSWTPDKVKTALLETADIAPKEIADIAYGAGRNNY 432
DB 241 AAGTSMATPHVAGIALLLQAHPSWTPDKVKTALLETADIAPKEIADIAYGAGRNNY 300
QY 433 KAIKYDDYAKLTFTGSAADKGSATHTFDVSGATFTATLTYMDTSSDIDLTYLDPNGNEV 492
DB 301 KAIYDNYAKLVFTGYVANKGSQTHQFVIGASFTVATLYMDNANSDDLTYLDPNGNOV 360
QY 493 DYSTAYYGFEEKYGVNPTAGTWTVKVSYKGAANYOVVSDGSLSSGSGGNPNPNP 552
DB 361 DYSTAYYGFEEKYGVNPTAGTWTIKVSYSGSANYOVVSDGSLSSGSGGNPNPNP 415
QY 553 NPFTTDTQTFTGSVNDYMDTSDFTTMVNSGATKTIGDLTFTDTSYNDLTYLDPNGNEV 612
DB 416 QPPTVADAKTFQKSDHYXDRSDFTFTVNSGATKLTGDLVFTDTSYNDLTYLDPNGNEV 475
QY 613 VDRSTSSNGEHEHYANPARGTWTFVYAYSTYGMADYOLKAVYTYG 659
DB 476 VDRSSEPSNIEHYEYLTLPAGTWTFVYAYSTYGMAYEYLTAKVYTG 522

RESULT 8
AAW94838
ID AAW94838 standard; protein; 522 AA.

XX AAW94838;

XX 06-MAY-1999 (first entry)

XX Hyperthermostable protease.

KW Hyperthermostable; protease; thermophilic; bacterium; subtilisin;
additive; drug; washing agent; foodstuff; chemical synthesis.

XX Pyrococcus furiosus.

XX Key Location/Qualifiers

XX Misc-difference 428 /label= Gly or Val

XX MO9856926-A1.

XX 17-DEC-1998.

XX 04-JUN-1998; 98MO-JP002465.

XX 10-JUN-1997; 97JP-00151969.

XX (TAKI) TAKARA SHUZO CO LTD.

XX Takakura H, Morishita M, Shimojo T, Asada K, Kato I;

XX WPI, 1999-080907/07.

PT Recombinant hyperthermostable protease from Pyrococcus furiosus - and
PT gene encoding it, for large scale production of the protease for
PT industrial use.

PS Claim 1; Page 39-41; 82pp; Japanese.

CC The invention relates to a hyperthermostable protease derived from a
CC thermophilic bacterium (especially Pyrococcus furiosus). The protease has
CC working temperature 40-110 deg.C (optimum 80-95 deg.C), working pH 5-10
CC (optimum 6-8), and retains more than 90% of its activity after 8 hours at
CC 95 deg.C. The invention also provides gene sequences encoding a
CC polypeptide of formula SIG-Ala-Gly-Gly-Aasn-PRO, where SIG is a signal
CC peptide from subtilisin, and PRO is the above protease. Host cells
CC (especially Bacillus strains) transformed with vectors comprising the
CC genes are used for the recombinant production of the protease. The
CC hyperthermostable protease which can be prepared in quantity suitable for
CC industrial use, can be used as an additive for drugs, washing agents and
CC foodstuffs and for chemical synthesis

SO Sequence 522 AA;

Query Match 68.5%; Score 2349.5; DB 2; Length 522;
Best Local Similarity 85.2%; Pred. No. 2e-140;
Matches 449; Conservative 22; Mismatches 51; Indels 5; Gaps 1;

QY 133 AELBGLDESSAAQWATYVMNLGYDGSGITIGITIDGIDASHPDLOGKVTGWDFNGRSY 192
DB 1 AELBGLDESSAAQWATYVMNLGYDGSGITIGITIDGIDASHPDLOGKVTGWDFNGRSY 60
QY 193 PYDDHGHGHTHVASIAAGTGAASNGKYKGMAPGAKIAGIKVLGADGSGSISTIIKGYEMAV 252
DB 61 PYDDHGHGHTHVASIAAGTGAASNGKYKGMAPGAKIAGIKVLGADGSGSISTIIKGYEMAV 120
QY 253 DNKDKYGIKVINLSLSSGSSDGTDSLISQAVNNAMDAGIYVVAAGNSGPNYTVGSPPA 312
DB 121 DNKDKYGIKVINLSLSSGSSDGTDSLISQAVNNAMDAGIYVVAAGNSGPNYTVGSPPA 180
QY 313 ASKYITVGAVDNSNNIASFSSRGPTADGRKPEVVAAGVUIIAPRASGTSMTGPIINDYTT 372
DB 181 ASKYITVGAVDKDVITFSFSSRGPTADGRKPEVVAAGVUIIAPRASGTSMTGPIINDYTT 240
QY 373 KASGTSMATPHVSGVGLILQAHPSWTPDKVKTALLETADIAPKEIADIAYGAGRNNY 432
DB 241 AAGTSMATPHVAGIALLLQAHPSWTPDKVKTALLETADIAPKEIADIAYGAGRNNY 300
QY 433 KAIKYDDYAKLTFTGSAADKGSATHTFDVSGATFTATLTYMDTSSDIDLTYLDPNGNEV 492
DB 301 KAIYDNYAKLVFTGYVANKGSQTHQFVIGASFTVATLYMDNANSDDLTYLDPNGNOV 360
QY 493 DYSTAYYGFEEKYGVNPTAGTWTVKVSYKGAANYOVVSDGSLSSGSGGNPNPNP 552
DB 361 DYSTAYYGFEEKYGVNPTAGTWTIKVSYSGSANYOVVSDGSLSSGSGGNPNPNP 415

```
QY 553 NPTPTDFTGTSVNDYMDTSDTFMTNVSNGATKITGDLTFPTSYNDLILYDPNGML 612
DB 416 QPEPTDARTFOKSDHYHYDRSDTFMTNVSNGATKITGDLVFPSTSHDDLILYDPNGKL 475
QY 613 VDRSTSSNSYEHVEYANPAAGTWTPLVYASTYGMADYOLKAVVYVG 659
DB 476 VDRSESPPNSYEHVEYLTLPAGTWTPLVYAYTYGMAYYELTAKVYVG 522

RESULT 9
AAM94836
ID AAM94836 standard; protein; 412 AA.
XX
AC AAM94836;
XX
DT 06-MAY-1999 (first entry)
XX
DE Hyperthermostable protease fragment.
XX
KM Hyperthermostable; protease; thermophilic; bacterium; subtilisin;
KW additive; drug; washing agent; foodstuff; chemical synthesis.
XX
OS Pyrococcus furiosus.
XX
PN MO9856926-A1.
XX
PD 17-DEC-1998.
XX
PF 04-JUN-1998; 98WO-JP002465.
XX
PR 10-JUN-1997; 97JP-00151969.
XX
PA (TAKI ) TAKARA SHUZO CO LTD.
XX
PI Takakura H, Morishita M, Shimojo T, Asada K, Kato I;
XX
DR WPI: 1999-080907/07.
XX
DR N-PSDB; AAX05920.
XX
PT Recombinant hyperthermostable protease from Pyrococcus furiosus - and
PT gene encoding it, for large scale production of the protease for
PT industrial use.
XX
PS Claim 2; Page 35-37; 82pp; Japanese.
XX
CC The invention relates to a hyperthermostable protease derived from a
CC thermophilic bacterium (especially Pyrococcus furiosus). The protease has
CC working temperature 40-110 deg.C (optimum 80-95 deg.C), working pH 5-10
CC (optimum 6-8), and retains more than 90% of its activity after 8 hours at
CC 95 deg.C. The invention also provides gene sequences encoding a
CC polypeptide of formula SIG-Ala-Gly-Gly-Asn-PRO, where SIG is a signal
CC peptide from subtilisin, and PRO is the above protease. Host cells
CC (especially Bacillus strains) transformed with vectors comprising the
CC genes are used for the recombinant production of the protease. The
CC hyperthermostable protease which can be prepared in quantity suitable for
CC industrial use, can be used as an additive for drugs, washing agents and
CC foodstuffs and for chemical synthesis
XX
SQ Sequence 412 AA;

Query Match 56.0%; Score 1918; DB 2; Length 412;
Best Local Similarity 89.5%; Pred. No. 3.3e-113;
Matches 367; Conservative 15; Mismatches 28; Indels 0; Gaps 0;

QY 133 AEELEGDESAQAQMATVYVNLGYDGSGITIGITDGTIDASHPLDQKVGWDFVNGRSY 192
DB 1 AEELEGDESAQAQMATVYVNLGYDGSGITIGITDGTIDASHPLDQKVGWDFVNGRSY 60
QY 193 PYDDHGHGTHVASIAAGTGAASNGKYKMAPGAKLAGIKVLGADSGSISTIIKGVEMAV 252
DB 61 PYDDHGHGTHVASIAAGTGAASNGKYKMAPGAKLAGIKVLGADSGSISTIIKGVEMAV 120
```

```
QY 253 DNKDKXGKIVNLSLSSQSSDGDTSLSQAVNNAMADAGIVCAAGNSGPNTYVSSPA 312
DB 121 DNKDKXGKIVNLSLSSQSSDGDTSLSQAVNNAMADAGIVVVAAGNSGPNKTYTSSPA 180
QY 313 ASKVITVGAVDSNDINASTSSRGPTADGRLKEEYVAPGVDDIIAPRASGTSKGTPIINDYT 372
DB 181 ASKVITVGAVDVKDYITSSSSRGPTADGRLKEEYVAPGVDDIIAPRASGTSKGTPIINDYT 240
QY 373 KASGTSMATPHYSGVAGLLOAHPSTPDKVTALILETADIYAPKEIADIAGAGRVNY 432
DB 241 AAPGTSMATPHVAGLIALLOAHPSTPDKVTALILETADIYAPKEIADIAGAGRVNY 300
QY 433 KAIXDDYAKLTFTGTSVADKGSATHTFDVSGATFVATLWMDTGSSDIDLILYDPNGNEY 492
DB 301 KAINVDNNAKLFTGTGVANKSGSQTHQFVLSGASFVATLYMDNANDDLILYDPNGNOY 360
QY 493 DYSYTAHYGFEKVGYYNPTAGTWTYKAVSYKGAANYQVDVSDGSLSSQG 542
DB 361 DYSYTAHYGFEKVGYYNPTDGTWTIKVSYSGSANYQVDVSDGSLSSQG 410

RESULT 10
AAR87009
ID AAR87009 standard; protein; 237 AA.
XX
AC AAR87009;
XX
DT 10-MAY-1996 (first entry)
XX
DE Hyperthermostable protease.
XX
KM Protease; hyperthermostable; thermostability.
XX
OS Pyrococcus furiosus.
XX
FH Key Location/Qualifiers
FT Misc-difference 196
FT note="unspecified amino acid"
XX
PN MO9534645-A1.
XX
PD 21-DEC-1995.
XX
PF 05-JUN-1995; 95WO-JP001095.
XX
PR 13-JUN-1994; 94JP-00130236.
XX
PR 26-JUL-1994; 94JP-00173912.
XX
PA (TAKI ) TAKARA SHUZO CO LTD.
XX
PI Mitra M, Yamamoto K, Morishita M, Asada K, Tsunabawa S, Kato I;
XX
DR WPI: 1996-049674/05.
XX
DR N-PSDB; AAT08131.
XX
PT Pyrococcus furiosus hyper:thermostable protease gene - useful for
PT recombinant prodn. of hyper:thermostable protease.
XX
PS Example 3; Page 70-71; 85pp; Japanese.
XX
CC The invention relates to the hyperthermostable protease of Pyrococcus
CC furiosus and its prodn. as a recombinant protein in transformants using a
CC vector carrying the protease gene (AAT08141). A genomic DNA sequence of
CC the invention is given in AAT08131 and its encoded protein in AAR87009
XX
SQ Sequence 237 AA;

Query Match 30.7%; Score 1051.5; DB 2; Length 237;
Best Local Similarity 87.9%; Pred. No. 1.3e-58;
Matches 203; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

QY 175 DLQKVIQWVDFVNGRSYPPYDDHGHGTHVASIAAGTGAASNGKYKMAPGAKLAGIKVLG 234
```


QY 165 IDTGIDASHEDLOGKVIIGWVDFVNGRSYPYDDHGHGTHVASIAGTGAASNGKYGKMGAPG 224
DB 61 IDTGIDASHEDLOGKVIIGWVDFVNGRSYPYDDHGHGTHVASIAGTGAASNGKYGKMGAPG 120
QY 225 AKIAGIKVIGADSGSISTTIKGVEMAVNDKDKYGIKVINLSIGSSQS 272
DB 121 AKIAGIKVIGADSGSISTTIKGVEMAVNDKDKYGIKVINLSIGSSQS 168

RESULT 13

AAR87011
ID AAR87011 standard; peptide; 188 AA.

AC AAR87011;

DT 10-MAY-1996 (first entry)

DE Peptide sequence.

KM Protease; hyperthermostable; thermostability.

OS Pyrococcus furiosus.

PN WO9534645-A1.

PD 21-DEC-1995.

PF 05-JUN-1995; 95WO-JP001095.

PR 13-JUN-1994; 94JP-00130236.

PR 26-JUL-1994; 94JP-00173912.

PA (TAKI) TAKARA SHUZO CO LTD.

PI Mita M, Yamamoto K, Morishita M, Asada K, Tsunashawa S, Kato I;

DR WPI; 1996-049674/05.

XX Pyrococcus furiosus hyperthermostable protease gene - useful for

PT recombinant prodn. of hyperthermostable protease.

PS Example 4; Page 66-67; 85pp; Japanese.

CC The invention relates to a gene (AAT08141) that codes for a

CC hyperthermostable protease (AAR87007) of Pyrococcus furiosus. 2 DNA

CC Sequence 188 AA: are provided encoding peptides (AAR87010-11)

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

ID AAW24126 standard; protein; 188 AA.
XX
XX AAW24126;
AC
XX
DT 17-OCT-2003 (revised)
DT 20-APR-1998 (first entry)
XX
DE Thermococcus protease fragment.

KM Protease; research reagent; thermal stability; thermococcus celer.

OS Thermococcus celer; DSM-2476.

PN WO9721823-A1.

PD 19-JUN-1997.

PF 07-NOV-1996; 96WO-JP003253.

PR 12-DEC-1995; 95JP-00323285.

PA (TAKI) TAKARA SHUZO CO LTD.

PI Takakura H, Morishita M, Yamamoto K, Mita M, Asada K;

PI Tsunashawa S, Kato I;

DR WPI; 1997-332794/30.

DR N-PSDB; AAT85676.

XX Protease(s) and genes encoding them obtained from Thermococcus and

PT Pyrococcus strains - have extremely high thermal stability and are useful

PT industrially and as research reagents.

XX Disclosure; Page 110-112; 159pp; Japanese.

PS This sequence represents a fragment of the protease from Thermococcus

CC celer DSM-2476 (see AAW24121 for full length sequence). This sequence is

CC a fragment of the protease of the invention. The proteases of the

CC invention have extremely high thermal stability. The proteases can be

CC used as research reagents, and industrially in the food, drug and

CC chemical industries. (Updated on 17-OCT-2003 to standardise OS field)

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

ID AAW13667 standard; protein; 734 AA.
XX
XX AAW13667;
AC
XX
DT 06-OCT-1997 (first entry)
DT

KM Protease; research reagent; thermal stability; thermococcus celer.

OS Thermococcus celer; DSM-2476.

PN WO9721823-A1.

PD 19-JUN-1997.

PF 07-NOV-1996; 96WO-JP003253.

PR 12-DEC-1995; 95JP-00323285.

PA (TAKI) TAKARA SHUZO CO LTD.

PI Takakura H, Morishita M, Yamamoto K, Mita M, Asada K;

PI Tsunashawa S, Kato I;

DR WPI; 1997-332794/30.

DR N-PSDB; AAT85676.

XX Protease(s) and genes encoding them obtained from Thermococcus and

PT Pyrococcus strains - have extremely high thermal stability and are useful

PT industrially and as research reagents.

XX Disclosure; Page 110-112; 159pp; Japanese.

PS This sequence represents a fragment of the protease from Thermococcus

CC celer DSM-2476 (see AAW24121 for full length sequence). This sequence is

CC a fragment of the protease of the invention. The proteases of the

CC invention have extremely high thermal stability. The proteases can be

CC used as research reagents, and industrially in the food, drug and

CC chemical industries. (Updated on 17-OCT-2003 to standardise OS field)

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DE Streptomyces viridoporus dhpa gene product.
 XX asymmetric hydrolyase; dhpa; 4-substituted-1,4-dihydropyridine;
 KW derivative; Streptomyces viridoporus; ester; chiral; synthesis;
 KM cardiovascular; treatment; hypertension; ischaemic heart disease.
 XX Streptomyces viridoporus.
 OS
 FH Key Location/Qualifiers
 FT Peptide 205..724
 FT /note="fragment of dhpa protein; see AAM13666"
 XX
 PN W09705243-A1.
 XX 13-FEB-1997.
 PD 13-FEB-1997.
 XX 30-JUL-1996; 96WO-JP002147.
 PF 31-JUL-1995; 95JP-00212975.
 PR 29-FEB-1996; 96JP-00067478.
 XX
 PA (SAOC) MERCIAN CORP.
 XX
 PI Aisawa A, Matsumaji M, Tsuruta T, Dobashi K, Nakashima T;
 PI Isehiki K, Yoshiooka T;
 DR WPI; 1997-145682/13.
 DR N-PSDB; AAT61454.
 XX
 PT Asymmetric hydrolyase gene derived from Streptomyces viridoporus - acts
 PT on 4-substituted-1,4-di:hydro:pyridine derivatives to produce chiral
 PT derivatives useful for synthesis of cardiovascular drugs.
 XX
 PS Claim 3; Page 49-55; 78pp; Japanese.
 XX
 CC This sequence is an asymmetric hydrolase encoded by the Streptomyces
 CC viridoporus dhpa gene. The enzyme acts on 4-substituted-1,4-
 CC dihydropyridine derivatives. The enzyme allows the efficient conversion
 CC of 4-substituted-1,4-dihydropyridine esters to chiral partially
 CC hydrolysed derivatives, for use in the synthesis of cardiovascular drugs
 CC suitable for the treatment of e.g. hypertension and ischaemic heart
 CC disease
 CC
 SQ Sequence 734 AA;
 Query Match 20.2%; Score 692.5; DB 2; Length 734;
 Best Local Similarity 33.6%; Pred. No. 3.3e-35;
 Matches 202; Conservative 62; Mismatches 223; Indels 115; Gaps 21;
 QY 110 LTGGKAKLSVRIQBDYKVTVAELGLDESAQVMATYVWNLGYDGSGITTIIDTGI 169
 DB 182 VTNGDRTASGIAHWLD-----GVRRAALDTSVGOIGAPKAMSAGYDGKVKIADVDTGV 236
 QY 170 DASHPDLOGKIVGWVFNNGRSYRYDDHGHGHTVASTAGTGAASNGKYGMAKGLAG 229
 DB 237 DTSHPDLOGKRVTAASKFTAPAG-AGDKVGHGHTVASTAGTGAQSKKYGVAFGAAILN 295
 QY 230 IKVLGADSGSISTTIKGVEMAVDNKKGIVKINLSGSOSSDGTDSLQAVNN-AMD 288
 DB 296 GKVLDDSGFGDDGILLAGWMAA-----AQGADVNNMSLGMDTPE-TDPLEAAVDKLSAE 350
 QY 289 AGIVCVAAAGNSGPNPTTYTGGSPAASKVITVGAVDSDNINASTSSRGP-TADGRLKEPV 347
 DB 351 KGVLFALIAAGNEPE--SIGSPGSAADALTVGAVDDKDLADFSSTGPRIGDAIKPDVT 408
 QY 348 APCVDIIAPASGTSMTPIIND---YTTKASGSMATPHVSGVGAALILQHPMTEDKTK 404
 DB 409 ARGVDTTAASABENDIGGVGEGPAGYMTISGTSMTFHVAGAAALLKQHPMTSABLK 468
 QY 405 TALIFTADIVAPREIADIVAGRVNVTAKI----- 436
 DB 469 GAL--TGSTKGK-YTPFEGSGRIQADKALQOTVLAIDPVSVFVGQWPHHTDDEPVTQK 525

QY 437 -----YDDVAKLTFTGSVADKGSATHTFDVSGATFTATLTLYMDTGSSDIDLILYDPV 488
 DB 526 LTRNLTQDVTYKLTSTATDPRKGAAPACFTLTGATVTPA---GGSASVDMTADTRL 582
 QY 489 GNEVDYSYTAAYYGFKEKVGYNPAGTWTYKVSYY-----KGAANYQDVVSDGS 537
 DB 583 GGTVDGAYSAYVVAATGGGQTVRTAAAVQREVESEYDVTVRHIGRDKPTTEHLTDLIGVAG 642
 QY 538 LSGSGGNRPNPNPPTPTDTQTF---TGS-VNDYMDTSPFTM-----NNSGA 585
 DB 643 LSGSGRYG-----APATDTATLRLPKGYTLVDSMTAKDFGLTKGGIDMLVQPKLSV 693
 QY 586 TKITGDLTPEPT-SYNDLIDLILYDPNGNLVDRSTSSSYEHVEYANPAPGTWTFVVAYST 644
 DB 694 TKDT-TLTIDARTTKAADITVPDPK-----AKPLSAT---IGTYDT 731
 QY 645 YG 646
 DB 732 AG 733

Search completed: January 6, 2007, 22:24:51
 Job time: 146 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2007 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2007, 22:25:12 ; Search time 27 Seconds
(without alignments)

2348.402 Million cell updates/sec

Title: US-10-800-684-5

Perfect score: 3428

Sequence: 1 MKGKALIVILVIGLVVGS.....YASTYGMADYQLKAVVYG 659

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------------|
| 1 | 688 | 20.1 | 442 | 2 | A69587 intracellular alka |
| 2 | 667.5 | 19.5 | 444 | 2 | B83891 intracellular alka |
| 3 | 555 | 16.2 | 382 | 1 | SUBSN subillisin (EC 3.4 |
| 4 | 554 | 16.2 | 379 | 1 | SUBSCL subillisin (EC 3.4 |
| 5 | 544 | 15.9 | 382 | 2 | T28159 subillisin (EC 3.4 |
| 6 | 540.5 | 15.8 | 1398 | 2 | pyrrolysin (EC 3.4 |
| 7 | 537.5 | 15.7 | 380 | 2 | A49778 high-alkaline seri |
| 8 | 535 | 15.6 | 274 | 1 | SUBSD subillisin (EC 3.4 |
| 9 | 524 | 15.3 | 381 | 2 | JH0778 subillisin (EC 3.4 |
| 10 | 523 | 15.3 | 275 | 2 | JC1085 subillisin (EC 3.4 |
| 11 | 523 | 15.3 | 381 | 1 | SUBSI subillisin (EC 3.4 |
| 12 | 520 | 15.2 | 381 | 1 | SUBSS subillisin (EC 3.4 |
| 13 | 520 | 15.2 | 381 | 2 | JQ1487 subillisin (EC 3.4 |
| 14 | 508 | 14.8 | 361 | 2 | A46373 high-alkaline seri |
| 15 | 508 | 14.8 | 361 | 2 | G83756 subillisin-type al |
| 16 | 505.5 | 14.7 | 806 | 2 | A41341 microbial serine p |
| 17 | 498.5 | 14.5 | 272 | 2 | A23624 subillisin (EC 3.4 |
| 18 | 496.5 | 14.5 | 384 | 2 | JC4802 alkaline proteinase |
| 19 | 496.5 | 14.5 | 757 | 2 | G84120 subillisin-type pr |
| 20 | 496 | 14.5 | 374 | 2 | I39781 subillisin (EC 3.4 |
| 21 | 491.5 | 14.3 | 401 | 2 | I39974 serine proteinase |
| 22 | 489.5 | 14.3 | 279 | 1 | SUMYTV thermolase (EC 3.4 |
| 23 | 488 | 14.2 | 799 | 2 | G83753 subillisin-type pr |
| 24 | 486.5 | 14.2 | 378 | 2 | A33973 high-alkaline seri |
| 25 | 460 | 13.4 | 715 | 2 | JC4908 alkaline serine pr |
| 26 | 452.5 | 13.2 | 1433 | 1 | A36734 bacillipeptidase F |
| 27 | 451.5 | 13.2 | 372 | 2 | D83735 subillisin-type al |
| 28 | 449 | 13.1 | 627 | 2 | D75393 serine proteinase, |
| 29 | 446.5 | 13.0 | 535 | 2 | B82358 alkaline serine pr |

| | | | | | |
|----|-------|------|-----|---|----------------------------|
| 30 | 443 | 12.9 | 440 | 2 | H72784 probable alkaline |
| 31 | 440 | 12.8 | 488 | 2 | A11930 proteinase (import |
| 32 | 439.5 | 12.8 | 519 | 2 | S71451 haloalysin R4 (EC 3 |
| 33 | 437.5 | 12.8 | 397 | 2 | JM0075 cysteine-dependent |
| 34 | 428.5 | 12.5 | 513 | 1 | A35742 agmatinase (EC 3.4 |
| 35 | 426.5 | 12.4 | 534 | 1 | US0173 alkaline proteinase |
| 36 | 423 | 12.3 | 588 | 2 | C83836 subillisin-type pr |
| 37 | 418 | 12.2 | 745 | 2 | JC6119 subillisin-type pr |
| 38 | 417.5 | 12.2 | 401 | 2 | A57690 aerolysin precuro |
| 39 | 417 | 12.2 | 747 | 2 | T06580 subillisin-like pr |
| 40 | 416 | 12.1 | 645 | 1 | SUBSMP serine proteinase |
| 41 | 409.5 | 11.9 | 525 | 2 | G84406 haloalysin (importe |
| 42 | 401.5 | 11.7 | 754 | 2 | T06579 subillisin-like pr |
| 43 | 401 | 11.7 | 401 | 2 | A42605 haloalysin (EC 3.4 |
| 44 | 400 | 11.7 | 319 | 2 | I39866 microbial serine p |
| 45 | 398 | 11.6 | 754 | 2 | T06577 subillisin-like pr |

ALIGNMENTS

Result 1

A69587 intracellular alkaline serine proteinase aprx - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C:Accession: A69587

R:Kunet, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertec

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabre, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funai, S.; Gallizzi, A.; Galler

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.;

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel,

Y, M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schoeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpeira, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K.

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Zumbstein, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98040403; PMID:9384377

A:Accession: A69587

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-442 <KUN>

A:Cross-references: UNIPROT:O31788; UNIPARC:UPI000006043F; GB:Z99113; GB:AL009126; NID:9;

A:Experimental source: strain 168

C:Genetics:

A:Gene: aprX

F:146-398/Domain: subillisin homology <SBT>

Query Match 20.1%; Score 688; DB 2; Length 442;

Best local similarity 39.7%; Pred. No. 1.8e-30; Mismatches 132; Indels 44; Gaps 14;

Matches 167; Conservative 78;

| | | |
|----|-----|---|
| QY | 43 | TFGLFRKI--QKLNNEISTIVIEENHREK--EIAVRVLEL-MGAKRYVYHIIPTA 96 |
| DB | 35 | TFCLFHKFFETQLQRKKKSVIIEFECHGTFQWAGVQLQEKRSKIKSRFNKNCSS 94 |
| QY | 97 | ADLKYRDLVLSGLTGKGRKSGVRFIOEDYVYVTSAELEGDESAQVM--ATVYVWNGS- 154 |
| DB | 95 | AE-----VTPSALHSLISECSNIR-----FVYANREVKALDTRTESHAKEVVRNQ 142 |
| QY | 155 | -YDGSGITIGLIDTSDASHPDLQGVIGWDFVNGRSYFYDDHGHGTHVASTAAGTGA 213 |
| DB | 143 | TLTGKGVYAVVDITG-YPHPDLBGRITIGFADMVNQKTEPFYDDNHGHTCAQDVASGAS 201 |
| QY | 214 | SNKTKYGNAPGAKLAGIKVLGADSGSISTTIKYEMAV---DNKDKYGVINISLGS 269 |
| DB | 202 | SSGOYRGAPAPFANLIGVVLNKGSGTLADIIEGVWCIOYVNEPDE-PIDIMWSLGS 260 |

```

Qy 270 S---OSSDGTSLQAVNNMADAGIVCVAAGSGPNTYTVGSPAAASKVITTVGAVDSN- 325
Db 261 DALRYDHEQEDPLVRAVEASAGIVVCVAAGSGPDSQITIASGVSEKVTITVALDDNN 320
Qy 326 -----DNIAFSSRGPTADRLKPEVVAPEVDIIAPASGT-----SMGTPINDYTTKA 374
Db 321 TASDDDDTVASFSSRGPTVYGEKPEVDIIAPGVNIIISLSPSYIDKLQKSRVSGVFTM 380
Qy 375 SGTSMATPHVSGVALLQAHPSWTPDKVTALLETADIYAPKEIADIYAGAGVWNYKA 434
Db 381 SGTSMATPHVSGVALLQAHPSWTPDKVTALLETADIYAPKEIADIYAGAGVWNYKA 438
Qy 435 I 435
Db 439 V 439

```

RESULT 2

```

B83891
Intracellular alkaline serine proteinase aprX [imported] - Bacillus halodurans (strain C
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: B83891
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: B83891
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-444 <SNO>
A/Cross-references: UNIPROT:Q9RBU7; UNIPARC:UPI00000C3D43; GB:AP001513; GB:BA000004; NIT
A/Experimental source: strain C-125
C/Genetics:
A/Gene: aprX

```

```

Query Match 19.5%; Score 667.5; DB 2; Length 444;
Best Local Similarity 36.7%; Pred. No. 2.3e-29;
Matches 156; Conservative 80; Mismatches 122; Indels 67; Gaps 12;

```

```

Qy 35 VERNKGLTTPGLFRKIQKLNPNNEISTVYFENHREKEIAVRVLELMGAKVRVYHIIIPA 94
Db 57 VEINOGCPDGG-----SMVYENHLKKGHK-----COLRHSFKIRC 92
Qy 95 IAADL---KVRDLVLSGLTGAKAKLSGVRFQSDYKVTVSALLEG-LDESAQVWATYV 150
Db 93 VSNATVPESLOEMLVCKDIR-----KIYLNREVALDTPAVESAQAPEV 136
Qy 151 WNLG--YDSGRTITGIIDPIDASHDLOGKVIGWDPFNGRSYPRDDHGHGTHVASTAA 208
Db 137 IRNGEITLTGKDVITIAVIDGI-YPHEDLEGRIKAFVDPVNGQREEPYDDNGHGTGCGDAA 195
Qy 209 GTGAASNGKYGKAPAKLAGIKVLGADSGSISTTIKGVEMAVDKDKY---GIKVINL 265
Db 196 GNGASDSGQYRGAPAPAVNYGVVKNQMGSLSEIMQGVEMQIQVNEHPDDPIHIIIM 255
Qy 266 SIQGS---OSSDGTSLQAVNNMADAGIVCVAAGSGPNTYTVGSPAAASKVITTVGAV 322
Db 256 SIQGQALPYENEGEDPMVRIVEEAMNAGITVCAAGNSGPDQITIASPVSSEKVTITVAL 315
Qy 323 D-----SNDNIAFSSRGPTADRLKPEVVAPEVDIIAPASGT-----SMGTPINDY 370
Db 316 DDRDITDRDDDDVAPSSSRGPTITYGKPEVDIIAPGVNIIISLSPSYIDKLQKSRVSG 375
Qy 371 YTKASGTSMATPHVSGVALLQAHPSWTPDKVTALLETADIYAPKEIADIYAGAGVWNY 430
Db 376 YTMMSGTSMATPHVSGVALLQAHPSWTPDKVTALLETADIYAPKEIADIYAGAGVWNY 433
Qy 431 VKKAI 435
Db 434 AEGAI 438

```

RESULT 3

SUBSN

subtilisin (EC 3.4.21.62) BPN' precursor - Bacillus amyloliquefaciens

N/A/Alternate names: subtilisin Novo

C/Species: Bacillus amyloliquefaciens

C/Date: 24-Apr-1984 #sequence_revision 28-Aug-1985 #text_change 05-Oct-2004

C/Accession: B25415; A93495; T4584; A92033; A00970

R/Vasantha, N.; Thompson, L.D.; Rhodes, C.; Banner, C.; Nagle, J.; Filipula, D.

J. Bacteriol. 159, 811-819, 1984

A/Title: Gene for alkaline protease and neutral protease from Bacillus amyloliquefaciens

A/Reference number: A25415; MUID:85006739; PMID:6090391

A/Accession: B25415

A/Molecule type: DNA

A/Residues: 1-382 <VAS>

A/Cross-references: UNIPROT:P00782; UNIPARC:UPI0000136180; GB:K02496; NID:g142525; PIDN:f

A/Experimental source: ATCC 23844

R/Wells, J.A.; Ferrari, E.; Henner, D.J.; Ercell, D.A.; Chen, E.Y.

Nucleic Acids Res. 11, 7911-7925, 1983

A/Title: Cloning, sequencing, and secretion of Bacillus amyloliquefaciens subtilisin in f

A/Reference number: A93495; MUID:84069812; PMID:6316278

A/Accession: A93495

A/Molecule type: DNA

A/Residues: 1-382 <WEL>

A/Cross-references: UNIPARC:UPI0000136180

A/Accession: T44584

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 'W', 8-382 <WEL>

A/Cross-references: UNIPARC:UPI0000136180; EMBL:X00165; NID:g39337; PIDN:CAA24990.1; PID

R/Marland, F.S.; Smith, E.L.

J. Biol. Chem. 242, 5198-5211, 1967

A/Title: Subtilisin BPN'. VII. Isolation of cyanogen bromide peptides and the complete an

A/Reference number: A92033; MUID:68086682; PMID:6065094

A/Accession: A92033

A/Molecule type: protein

A/Residues: 108-162, 'PV', 165-167, 'D', 169-194, 'SA', 197-204, 'DA', 207-264, 'ST', 267-357, 'Q', ;

A/Cross-references: UNIPARC:UPI0000112C9B

R/Kraut, J.

In The Enzymes, 3rd ed., vol.3, Boyer, P.D., ed., pp.547-560, Academic Press, New York, 1

A/Title: Subtilisin, X-ray structure.

A/Reference number: A94443

A/Comment: annotation; X-ray crystallography, 2.5 angstroms; active site

C/Comment: Secretion of subtilisin is associated with the onset of sporulation, and many

not necessary for normal sporulation.

C/Genetics:

A/Start codon: GTG

C/Superfamily: Subtilisin, subtilisin homology

C/Keywords: hydrolase; serine proteinase

F/1-32/Domain: signal sequence #status predicted <SIG>

F/133-107/Domain: activation peptide #status predicted <APT>

F/108-382/Product: subtilisin BPN' #status experimental <MPT>

F/130-342/Domain: subtilisin homology <SPT>

F/139,171,328/Active site: Asp, His, Ser #status experimental

```

Query Match 16.2%; Score 555; DB 1; Length 382;
Best Local Similarity 35.1%; Pred. No. 2.7e-23;
Matches 158; Conservative 65; Mismatches 141; Indels 86; Gaps 16;

```

```

Qy 1 MKGKALVLIVLIGV---VGSVAAPREKVGQVRNVEKNGYLLPGLFRKIQKLNPN 56
Db 1 MRGKVMVLSLFPALALITFMAFGSTSSA---QAAGKSGEKY----- 40
Qy 57 EISTVYFENHREKEIAVR---VLELMGAKVRVYHIIIPAIADLK---VRDLVLSGL 110
Db 41 -----IYGFQKMTSMGAKKQDVISERKGVQVQPFKVVDAASATLNEKAVKEL----- 89
Qy 111 TGGAKAKLSGVRFIEDYKVTVSALLEGIDESAAQVMAVYWNILGDSGRTITGIIDTGD 170
Db 90 ----KQDPVAVYVEDH--VAHAAYAGVPYGVSGQIKAPALHSQGYTSNVAIVDSDIG 143
Qy 171 ASHPDLOGKVIGWDPFNGRSYPR-DDHGHGTHVASTAAGTGAASNGK--YKGAAPAKL 227
Db 144 SSHPDL--KVAAGASMPVSETPNPQDNNSHGTHV---ACTVAALNNSIGVLGVAAPSASL 197

```

QY 228 AGIVKLAGDSSGSI::STIIKGEVAVDNDKKDGKIVINL:SLGSSQGSPOGDSLSQAVNNAM 287

Db 198 YAVVVLKLGDSGGQYSWIIINGIEMAIAN---NMVDYINSLG---GPSGSALLKAAVDAAV 250

QY 288 DAGIVVCVAANGSP--NTYTVGSPAAASKVITYGAVDSDNDNIASFSSRGFTADGRLKE 345

Db 251 ASGVVVVAAAAGNEGTSGSSSTVGVPGKYPVIAVGAVDSSNQGRASFSSVGPELD----- 304

QY 346 VVAGGVDDIAPRASGTSKGPINDIYTYKASGTSMAATPHVSGVALLLOAHSSWTPDKYKT 405

Db 305 VMAEGVSI-----QSTLPGNKYGAYNGTSMASPHVAGAAALLISKPHNMNTQVRS 355

QY 406 ALIETADIVAPKEIAD-IAYGARGVNYKA 434

Db 356 SLENTT-----TKLGDSFYTGKGLINVOAA 380

RESULT 4

subtilisin (EC3.4.21.62) Carlsberg precursor - *Bacillus licheniformis*
C:Species: *Bacillus licheniformis*
C:Date: 30-Jun-1998 #sequence_revision 30-Jun-1998 #text_change 05-Oct-2004
C:Accession: A24111; A00968
R:Jacobs, M.; Eliasson, M.; Uhlen, M.; Flock, J.I.
Nucleic Acids Res. 13, 8913-8926, 1985
A>Title: Cloning, sequencing and expression of subtilisin Carlsberg from *Bacillus licheniformis*
A:Reference number: A24111; MUID:86093688; PMID:3001653
A:Accession: A24111
A:Molecule type: DNA
A:Residues: 1-379 <JAC>
A:Cross-references: UNIPROT:P00780; UNIPARC:PI00000136181; GB:X03341; NID:g487721; PIDN:R.581
A:Experimental source: strain NCIB5816
J:Smith, E.L.; Delange, R.U.; Evans, W.H.; Landon, M.; Markland, F.S.
J. Biol. Chem. 243, 2184-2191, 1968
A>Title: Subtilisin Carlsberg. V. The complete sequence; comparison with subtilisin BPN'
A:Reference number: A00968; MUID:68234702; PMID:4967581
A:Accession: A00968
A:Molecule type: Protein
A:Residues: 106-206, 'S', 208-232, 'A', 234-261, 'N', 263-264, 'S', 266-315, 'N', 317-379 <SMT>
A:Cross-references: UNIPARC:UPI0000020021
C:Comment: Secretion of subtilisin is associated with the onset of sporulation, and many
not necessary for normal sporulation.
C:Superfamily: Subtilisin; subtilisin homology
C:Keywords: extracellular protein; hydrolase; serine proteinase
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-105/Domain: propetide #status predicted <PT>
F:106-379/Product: subtilisin Carlsberg #status experimental <MPR>
F:128-339/Domain: subtilisin homology <SMT>
F:137-168, 325/Active site: Asp, His, Ser #status predicted

```

Query March 16.2% Score 554; DB 1; Length 379;
Best Local Similarity 34.1%; Pred. No. 3.1e-23;
Matches 149; Conservative 65; Mismatches 149; Indels 74; Gaps 14;

QY 3 GLKALLIVILVLGVGSVAAAPEREKVEQVNVNEKNTGLTLPGLEFRKIQLKLPNEBISTV 62
      :::::
Db 10 GMLTAFMLVFLPMARSDSASAQPAK-----NVEEDY-----I 41

QY 63 IVEPNH-REKEIARVLELMGAKVRYVYHIIIPALADLKVRDLVLSGTGKAKLSGR 121
      :::::
Db 42 VGFPSGVATSVKDDIIKESGKVDKOPRIINAARAKAKIDKEALKEVK-----NDPVA 94

QY 122 FIODYKVTVAELEGDELSAAQVMATVVMNLGYDGSITGIIIDTGDIDASHPLDQKVI 181
      :::::
Db 95 IYBEDH--VAALAAQTVRYGILPKADKQVQAQGFGRANVKVAVLDTGIQASHPLD--NVV 150

QY 182 GWVDFVNGRSYYPYDDHGHGTHVASTAAGTGAASNGK--YKMAPGAKIAGIKVLGADSG 239
      :::::
Db 151 GGAASFVAGEAAYNTDGNHGTHV---AGTVAALDNTTGLVLGAAPSVSILYAVKVLNNSGG 206

QY 240 SIIITIKVEVAVDNKKYGIKVINLISGGSSSDGTDLSQAVNNMADAIIVCVAAAG 299
      :::::

```

[illegible]

RESULT 5

Subtilisin (EC 3.4.21.62) Sendai precursor - Bacillus sp.
 C:Species: Bacillus sp.
 C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #ext_change 05-Oct-2004
 C:Accession: J139780
 R:Yamagata, Y.; Ieshiki, K.; Ichishima, E.
 Enzyme Microb. Technol. 17, 653-663, 1995
 A:Title: Subtilisin Sendai from alkalophilic Bacillus sp.: molecular and enzymatic properties
 A:Reference number: J139780; MUID:95329264; PMID:7605625
 A:Accession: J139780
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-182 <RES>
 A:Cross-references: UNIPROT:Q45522; UNIPARC:UPI00000855E9; GB:D29688; NID:g995963; PTDN:E
 C:Genetics:
 A:Gene: aprQ
 A:Start codon: TTG
 C:Superfamily: Subtilisin, subtilisin homology
 C:Keywords: hydrolase; serine proteinase
 ;136-342/Domain: subtilisin homology <SBT>

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 15.9% | Score 544; | DB 2; | Length 382; |
| Best Local Similarity | 33.6% | Pred. NO. 1.1e-22; | | |
| Matches 150; | Conservative 69; | Mismatches 131; | Indels 96; | Gaps 17 |

```

Qy 10 VILVIGLWVGSVAAPAEKK-----VEORANEKKNVGLTLPLFRKIQKLPNE 57
Db 14 LILSILSLTSTVSABEOKKQYLIGENOLQTFEVESDKQ-----55
Qy 58 ELSTVLPFNHEKELEIARVLEIMGAKVRYYHIIIPALMDK--VRDLVVISGLTGK 114
Db 56 --SEMSLFAENVDESIEMEL-----YEFEDIPVSVSELPEDVYDL-----95
Qy 115 AKISGRFIFQEDKYTVASALE--GLDESSAQWATYMNIGYSGGTTGIDTGIDAS 172
Db 96 EKDPSTTYEBEDIEVITNQVTPMEI---TRVQAPMTAGYGTGVRVAVLTGT-ST 15G
Qy 173 HPDLQSKVIGWDFVNGRSYPYDDHGHTVHASIAGTGAASNGK--YKMAFGAKAGI 230
Db 151 HPPL--NINGGVSFPGEPSTYDGGHGHTV---AGTIALNNSIGVGCAPPAELIYAV 204
Qy 231 KVLGADSGSISTIIKGYEAVDNDRKXYGIKVINLSLSSGSSDGTSLSOAVNAMDAG 290
Db 205 KVLGANGSSVSASIIAQGLQMTAQN---NIHVANISLGSFV--GSOTLLEIYAVQATNAG 257
Qy 291 IYVCVAANGSGNWTYVSRAASKVITVGAVDSDNIIASFSSGRPTADGLKREYVAPG 350
Db 258 VLVVAATGNNGSG--TVSTPARYAALVGTADONNPAASQYGTGLN-----IYAPG 308
Qy 351 VDLIAPRASGTSMGTPINDYTKTASGTSMTATPHVSGVALLILOAHPSWTPDKVTALLET 410
Db 310 VGI-----QSTPGNRYASLSGTSMTATPHVAGVAAIYKQKNPMSWTOIRKHLTST 360
Qy 411 ADIVAPKETADIANYGGRVNYKATK 436
Db 361 ATSLGNSN---QPSGLVNAEAAAR 382

```

RESULT 6
T28159
Pyrolysin (EC 3.4.-.-) - Pyrococcus furiosus
C/Species: Pyrococcus furiosus
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T28159
R/Voorhorst, W.G.B.; Eggen, R.I.L.; Geerling, A.C.M.; Platteeuw, C.; Slezan, R.J.; Vos, J. Biol. Chem. 271, 20426-20431, 1996
A/Title: Isolation and characterization of the hyperthermostable serine protease, pyrolysin, from *Pyrococcus furiosus*
A/Reference number: Z20481; MUID:96355370; PMID:8702780
A/Accession: T28159
A/Status: preliminary; translated from GB/EMBL/DBD
A/Molecule type: DNA
A/Residues: 1-1398 <VOO>
A/Cross-references: UNIPROT:P72186; UNIPARC:UPI000016FD59; EMBL:U55835; NID:G1556462; PI
A/Experimental source: DSM3638
C/Genetics:
A/Gene: pls
C/Keywords: hydrolase; serine proteinase

Query Match 15.8%; Score 540.5; DB 2; Length 1398;
Best Local Similarity 22.5%; Pred. No. 9.9e-22;
Matches 238; Conservative 103; Mismatches 251; Indels 465; Gaps 38;

QY 2 KGLKALLVILVIGLV--VGSVAAPAEKVEQVRNVEKNYGLLT---PGLFRKIQKN 54
DB 4 KGLTLVFLIALMLLSVVFHFVSAGTPVSESTSTILPMQVTVKEVSOALNALMKQG 63
QY 55 PNEISIVIVFENHREKEIAVRVLELGMKAVRVYHIIPIAIDL---KYRDLVLSGL- 110
DB 64 PN---MWLIITKEGKLEBEKTELEKAGAILDBNRLLMLLYKIKPEKKEINYSISLE 120
QY 111 ---TGKAKLSGVRPIQEDYKVTVSABELGDSAAQVM--ATYVMNLGDSGSGITGII 165
DB 121 KAMILNREVKLS-PIYVKQYK-TKEPELBPKNMSTWVINALQFIQEFQDSSGVVAVL 178
QY 166 DTGIDASH---PDLQGVYGVMP---VNGR----- 190
DB 179 DTGVDPNHPPLSTTPDGRKRIEMKQFTDGFVDTSPSPKVVNCTLIINTFQVAGLT 238
QY 191 ----- 190
DB 239 LNESTGLMEVYKTVVYVSNVTIGNITSANGIYHFGILPERYFPLFDGQEDFPYVLVN 298
QY 191 ----- 194
DB 299 STNGGYDIAVVDLDDYDFTDEVLQGVYVTVDAVFSTYYGFLVLAIDPNRGYAVF 358
QY 195 ---DDHGHTTVASIAAGTGAAS-----NGKY-----KGMAPG 224
DB 359 GMDGHGHGHTVAGTVAGYDSNDNDAMDWLSMTSGMEVFSRLYGMWDTYNTTDTVQGVARG 418
QY 225 AKLAGIKVLGADSGSISTILKGVENAVNKKYGIKVINLSL-GSSQSSDPTGDSLSQAV 283
DB 419 AQQMAIRVLRSDGRSGMMWIIISGMTAA---THGADVISMLGSGNAFLDGTDESVAV 474
QY 284 NNAAMD-GLVVCYCAANSQPNYTVGSSPAASKVITTVGAVD----- 323
DB 475 DELTEKYGVVPIAANBGPINIVGSPGATKATIVGAAAPINVGYYVSOALGYPDY 534
QY 324 -----SNDNIASSRSGPTADGRLKPEVVAPOVDIIAPASGTSKMTPIINDYTTAAS 375
DB 535 GYVYPAATYVRIALFFSSRGPRIIDGRIKPVVAPGIGY---SSLPMMIGADR---MS 587
QY 376 GTSMATPHVSGVALLIQA-HPS---WTPDKYKTLIETA-----DIYAPKEIADIAYGA 426
DB 588 GTSMATPHVSGVALLISGPKEGITYNDIILKIVLESATMLBEDPTTGCKTLELDGH 647
QY 427 GRVNV-----YKAT-----KYDYAKL----- 443
DB 648 GLAVNTYKSWELLKALINGTTLPIVDHMADKSYSDFAEYLGVDVIRGLVARNSSIPDIVEMHI 707

QY 444 -----TF-----TGSV----- 449
DB 708 KYVGDTEYRTFEIYATEPWIKPFVSGSVLENNTEFVLKVKYDVEGLEGLYVGRITIDD 767
QY 450 -----ADKSATHT-FVSGATFYT-----ATLYW 473
DB 768 PTPVIEDELINTVIAPEKFTPENNNTLMTWIDINGEVMTHFFTVBQVDVLYXANTTW 827
QY 474 DTGSSSIDLYIDPNC-----NEVDYSYAYVGFKEKVGYNPTAGTWTKVVSYGKANY 528
DB 828 DVG-----LYRPDMFVFPPYQLDYLPA-----VSNPMFGWELVMTGFNPAPLY 872
QY 529 QVDVVSQSLSSQSGGNPNPNPNPTPTDTQTFTGSVN-DYMPDSTFTMNV----- 582
DB 873 E-----SGFLVLIHGVETPS-----WYINRITLYDITNTEFSIEFNTIY 913
QY 583 -----SGATKITGDLTF-----DTSYNDLDLYL 606
DB 914 APIATLIPIGLGTYNASVESVGDGEFFIKGIEPEGTAEIKIRIGNPSVPSNDLDLYL 973
QY 607 DPNGLVDRSTSSNSENHVEYANPARGTWTFLVAYS 643
DB 974 DSKGNLVALDGNPTAEERVVEYPRGVYSIVHGS 1010

RESULT 7
A49778
high-alkaline serine proteinase (EC 3.4.21.-) precursor - *Bacillus alcalophilus* (strain
N/Alternate names: subtilisin homolog, high-alkaline
C/Species: *Bacillus alcalophilus*
C/Date: 12-Mar-1994 #sequence_revision 24-Feb-1995 #text_change 05-Oct-2004
C/Accession: A49778; J01244
R/van der Laan, J.C.; Gerltzse, G.; Mullerers, L.J.S.M.; van der Hoek, R.A.C.; Quax, W.J.
Appl. Environ. Microbiol. 57, 901-909, 1991
A/Title: Cloning, characterization, and multiple chromosomal integration of a *Bacillus a*
A/Reference number: A49778; MUID:91282483; PMID:2059048
A/Accession: A49778
A/Molecule type: DNA
A/Residues: 1-380 <VAN>
A/Cross-references: UNIPROT:P27693; UNIPARC:UPI000020CF5A; GB:D13157; NID:G216231; PIDN:BAA02442.1; PID:G
A/Experimental source: strain P27693, ATCC 31408
A/Note: amino end of mature protein confirmed by peptide sequencing
R/Itakami, H.; Kobayashi, T.; Kobayashi, M.; Yamamoto, M.; Nakamura, S.; Aono, R.; Horikoi
Biosci. Biotechnol. Biochem. 56, 1455-1460, 1992
A/Title: Molecular cloning, nucleotide sequence, and expression of the structural gene fo
A/Reference number: J01244; MUID:99043753; PMID:1368952
A/Accession: J01244
A/Molecule type: DNA
A/Residues: 1-195, S', 197-380 <TKA>
A/Cross-references: UNIPARC:UPI000020CF5A; GB:D13157; NID:G216231; PIDN:BAA02442.1; PID:G
A/Experimental source: *Bacillus alcalophilus* Vedder, ATCC 21522 (*Bacillus* sp. 221)
C/Keywords: hydrolase; serine proteinase; zymogen
F.1-27/Domains: signal sequence #status predicted <SIG>
F.128-111/Domains: activation peptide #status predicted <PRO>
F.112-380/Product: alkaline serine proteinase #status predicted <MAT>
F.113-340/Domains: subtilisin homology <SBT>
F.143-173/Active site: Asp, His, Ser #status predicted

Query Match 15.7%; Score 537.5; DB 2; Length 380;
Best Local Similarity 33.6%; Pred. No. 2.5e-22;
Matches 147; Conservative 74; Mismatches 140; Indels 77; Gaps 17;

QY 14 LGLVVGSVA-----AAPEKVGQVRNVEKNYGLTGPGRKIQKNPNEIS 60
DB 5 LGKVYASTALLISVAFSSSTASAAEAKKYLIGFNEQ-----EAVSEFVEQVANDAVA 59
QY 61 TVIVFENHREKEIAVRVLELGMKAVRVYHIIPIAIDLKYRDLVLSGLTGKAKLSGV 120
DB 60 IL-----SEEEVEIETL-----HEFTIPVLSTVELSPED---VDALDELPA----- 199
QY 121 RFIQEDYKVTVSABELGDSAAQVMATYVMNLGYDSGSGITGIIPTGIDASHPDQCKY 180

Db 100 STIEDEAVTTMA--OSVWGISRVQAPAAHNGLTGSGVAVLDTGI-STHPDL--NI 154

Qy 181 IGWDFVNGRSYPYDDHGHGTHVASTAGTGAASNGK--YKGAAPGAKLAGIKVLGADGS 238

Db 155 RGASAVPPEPSTODONGHGHV---AGTIALNNISIVLGAAPAAELIYAKVLGASGS 210

Qy 239 GSISTTIKVENAVDNKDKYKIKVNLISGSSGSDGTDLSQAVNNAMDAIVCVAAAG 298

Db 211 GSVSSIAQGLEWAGNN---GHHVANTLSGSPSPS---ATLEQAVNSATSRGLVVAASG 263

Qy 299 NSGPNNTYVTSQPAASKVITGAVDSNDNIASFSSGPTADGLKEVAVAPGDIAPRA 358

Db 264 NSGAG--SISYPRAYANAAVGTDDNNRPASFSQYAGLD-----IYAPGVNVQSTYP 315

Qy 359 SGTSMGTPINDYYTKASGTSMATPHYSGVAGALLIQAHPMTDPDKVTALLETADIAPKE 418

Db 316 GST-----YASINGTSMATPHYAGAAALVKKONPMSNVQIFNHLKATRLTSGSTN 366

Qy 419 IADIVAGRVVYKAIK 436

Db 367 L-----YSGGLVNAEAATR 380

RESULT 8

SUBSD
subtilisin (EC 3.4.21.62) DY - Bacillus subtilis (strain DY)

N:Alternate names: alkaline serine proteinase

C:Species: Bacillus subtilis

A:Variety: strain DY

C:date: 20-Sep-1984 #sequence_revision 20-Sep-1984 #text_change 05-Oct-2004

C:Accession: A00969; S02492

R:Nedkov, P.; Oberthur, W.; Brantlitzer, G.

Biol. Chem. Hoppe-Seyler 366, 421-430, 1985

A:title: Determination of the complete amino-acid sequence of subtilisin DY and its comp

A:Reference number: A00969; PMID:85279896; PMID:3927935

A:Accession: A00969

A:Molecule type: protein

A:Residues: 1-274 <MED>

A:Cross-references: UNIPARC:UPI000002D020

A:Experimental source: strain DY

R:Lilova, A.; Kleinschmidt, T.; Nedkov, P.

Biol. Chem. Hoppe-Seyler 368, 1479-1487, 1987

A:title: Reductive alkylation of lysine residues in subtilisin DY.

A:Reference number: S02492; PMID:88134577; PMID:3124865

A:Accession: S02492

A:Molecule type: protein

A:Residues: 1-212-12-24-27-29-43-45-93-95-123-135-137-140-142-169-171-183-187-221-23

A:Cross-references: UNIPARC:UPI0000055EF8; UNIPARC:UPI0000055EF8; UNIPARC:UPI000016B38C; BFA; UNIPARC:UPI0000172BFC; UNIPARC:UPI0000172BFC; UNIPARC:UPI0000172BFC; UNIPARC:UPI000

C:Comment: Secretion of subtilisin is associated with the onset of sporulation, and many

not necessary for normal sporulation.

C:Superfamily: Subtilisin; subtilisin homology

C:Keywords: extracellular protein; hydrolase; serine proteinase

F:23-234/Domain: subtilisin homology <SBT>

F:32,63,220/Active site: Asp, His, Ser #status predicted

Query Match 15.6%; Score 535; DB 1; Length 274;

Best Local Similarity 44.9%; Pred. No. 2-2e-22;

Matches 132; Conservative 36; Mismatches 90; Indels 36; Gaps 9;

Db 145 WATYVWNLGYGSGTIGTIDGIDASHPDLOGKVIYWDVFNHGSYPYDDHGHGTHVA 204

Db 11 IRADKYQAOGYKQAVKVGIIIDYIGVASHIDL--KVVGASFSVSGSYNTDNGHGHV- 67

Qy 205 SIAGTGAASNGK--YKGAAPGAKLAGIKVLGADSGSISTTIKVENAVDNKDKYKIRY 262

Db 68 ---AGVVAALDNTVTGLVAVPNSVLAIVKIVNSGSGTSAIVSGIEMATQN---GLDY 120

Qy 263 INSLSSGSGSDGTDLSQAVNNAMDAIVCVAAENSGP--NTYVSGPAAASKITYTG 320

Db 121 INMSLG---GPSGSLTKQAVDKAVASGIYVVAANNSGSGSQNTIGYPAKDYIAAG 177

Qy 321 AVDSNNINASFSSRGTADGRLEKPEVVAPOVDIIAPRASGTSMGTPINDYYTKASGTMA 380

Db 178 AVDSNNINASFSSVQ-----AELEVMAPGVSVSTYSNT-----YISLNGTMA 222

Qy 381 TPHVSGVAGALLIQAHPSTPDKVTALLETADIAPKEIADIYAGRVVYKAI 434

Db 223 SPHVGAALLISKPTLSASQVRRRLSTATNLDG----SPYQKGLINVAA 272

RESULT 9

JH0778

subtilisin (EC 3.4.21.62) NAT precursor - Bacillus subtilis (strain natto NC2-1)

N:Alternate names: natto proteinase; nattokinase; subtilisin Bsp

C:Species: Bacillus subtilis

C:date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Oct-2004

C:Accession: JH0778; J50601; J50517; JC2036

R:Nakamura, T.; Yamagata, Y.; Ichishima, E.

Biosci. Biotechnol. Biochem. 56, 1869-1871, 1992

A:title: Nucleotide sequence of the subtilisin NAT gene, aprN, of Bacillus subtilis (natt

A:Reference number: JH0778; PMID:93113095; PMID:1365081

A:Accession: JH0778

A:Molecule type: DNA

A:Residues: 1-381 <NAK>

A:Cross-references: UNIPROT:P35835; UNIPARC:UPI0000000AB8; GB:D25319; NID:G435439; PIDN:E

R:Sumi, H.

Kagaku To Seibutsu 29, 119-123, 1991

A:title: Natto kinase and fibrinolysis.

A:Reference number: J50601

A:Accession: J50601

A:Molecule type: protein

A:Residues: 107-381 <SUM>

A:Cross-references: UNIPARC:UPI00001565DB

R:Sumi, H.; Nakajima, N.

Nippon Nogelkagaku Kaishi 65, 1125-1127, 1991

A:title: Studies on fibrinolysis enzymes in fermentation food.

A:Reference number: J50517

A:Accession: J50517

A:Molecule type: protein

A:Residues: 107-381 <SU2>

A:Cross-references: UNIPARC:UPI00001565DB

R:Fujita, M.; Nomura, K.; Hong, K.; Ito, Y.; Asada, A.; Nishimuro, S.

Biochem. Biophys. Res. Commun. 197, 1340-1347, 1993

A:title: Purification and characterization of a strong fibrinolytic enzyme (nattokinase)

A:Reference number: JC2036; PMID:94107337; PMID:8280151

A:Accession: JC2036

A:Molecule type: protein

A:Residues: 107-381 <FNU>

A:Cross-references: UNIPARC:UPI00001565DB

C:Genetics:

A:Gene: aprN

A:Start codon: GTG

C:Superfamily: Subtilisin; subtilisin homology

C:Keywords: hydrolase; serine proteinase; zymogen

F:1-29/Domain: signal sequence #status predicted <SIG>

F:30-106/Domain: activation peptide #status predicted <PRO>

F:107-381/Product: subtilisin NAT #status experimental <MAT>

F:129-341/Domain: subtilisin homology <SBT>

F:138,139,170,327/Active site: Asp, Ser, His, Ser #status predicted

Query Match 15.3%; Score 524; DB 2; Length 381;

Best Local Similarity 37.9%; Pred. No. 1.4e-21;

Matches 145; Conservative 53; Mismatches 127; Indels 58; Gaps 15;

Qy 61 TYIVENHREKIANRVLEMGAKRYVYHIIIPALADL--KVPDLVLSGLTGKAKL 117

Db 46 TNSAMSAKKXQ---VISEKGVQKQFKYVNAATAUDEKAVLEL-----KQD 91

Qy 118 SGVRFQEDYKTVTAAL--LEGDSQAQVMATYVWNLGYDGGITIGTIDGIDASHPD 176

Db 92 PSVAIVAEEDH---IAHEVAQSVPIGISOIKAPALNSQGYTSNVAVAVIDSIDSHPD 148

Qy 177 QGKVTGWDFVNGRSYPYDD--HGHTHVASTAGTGAASNGK--YKGAAPGAKLAGIKVL 233

Db 149 --NVGASFSVPSFTNPYQDGSSTHGHV---AGTIALNNISIVLGVAVPSASLAVAKVL 202

QY 234 GADGSGISTIKGVEMAVNDKRYGKIVNLISGSSQSDGTDSLQAVNNAMDAGIV 293
 Db 203 DSGSGGGYVWINGIEIAIEN---NMDVIMSG---GPTGSLAKTVVDKAVSSGIV 255
 QY 294 CVAAGNSGP--NTYVGSPPAAKRVITVGVVDSNDNIASRSSGPTADGLKPEVAVPV 351
 Db 256 AAAAGNMGSSGSTSTVGVPAKYSTIIVGVVNSNCRASFSSVSGSEID-----VMAFGV 309
 QY 352 DIAPRASGSTMGPINDYTTKASGTMATPHVSGVALLQAHPSWTPDKXTALIEFA 411
 Db 310 SIOSTLPFGT-----YGANNGSMATPHVAGAAALILSKHPTMTNAQVRELRLESTA 360
 QY 412 DIVAPKEIADIYAGAGRVNYKA 434
 Db 361 TYLG---NSFYGKGLINVOAA 379
 RESULT 10
 JCI085
 subtlisin (EC 3.4.21.62) precursor - *Bacillus licheniformis*
 N/Alternate names: alkaline proteinase
 C/Species: *Bacillus licheniformis*
 C/Date: 09-Aug-1995 #sequence, revision 19-Oct-1995 #text_change 05-Oct-2004
 C/Accession: JCI085
 R/Lei, H.; Hong, Y.; Zhang, Y.Y.; Shen, T.J.
 Chinese Biochem. J. 9, 441-447, 1993
 A/Title: PCR amplifying, cloning and sequencing of the coding sequence of the alkaline F
 A/Reference number: JCI085
 A/Accession: JCI085
 A/Molecule type: DNA
 A/Residues: 1-275 <LEI>
 A/Cross-references: UNIPROT:Q9F7C2; UNIPROT:Q53521; UNIPROT:Q9FDF4; UNIPROT:Q9FDF2; UNIF
 A/Note: The translation of the start codon ATG is not given in this paper
 C/Suprafamily: Subtlisin, subtlisin homology
 C/Keywords: hydrolase; serine proteinase
 F/24-235/Domain: subtlisin homology <Str>
 F/33,64,221/Active site: Asp, His, Ser #status predicted
 Query Match 15.34; Score 523; DB 2; Length 275;
 Best Local Similarity 43.5%; Pred. No. 9.9e-22;
 Matches 128; Conservative 39; Mismatches 91; Indels 36; Gaps 10;
 QY 145 VMAVYVNLGYDSGITIGIITDIDGASHPDLOGKVIQWDFVNGRSYPYDHDGHTVA 204
 Db 12 IKADKVOAGQFKGAVVAVLDIGIQASHPD--NVVGGASFVAGAYNTDGHGHV-- 68
 QY 205 SIAGTGAASNGK--YKGAAPGAKLAGIKVLGADGSGSITTIKGVEMAVNDKOKTGIKY 262
 Db 69 ---AGTVAAADNTGTGVGVSVSLVAVKVLNSGSGSYSGIYGIEMATTN---GMDV 121
 QY 263 INTLSGSSQSDGTDSLQAVNNAMDAGIVCVAAGNSGP--NTYVGSPPAAKRVITV 320
 Db 122 INMSLG---GASGTAMKQAVNDIYARGVVAAAAGNSGSGNTITIGYPAKIDSIANG 178
 QY 321 AVDSNDNIASFSSRGPTADRLKPEVAVPGVDIIAPRASGTMGPINDYTTKASGTSMA 360
 Db 179 AVDSNENRASFSSVSG---AELEVWAPG-----AGVSTYPTNTVAT--LNGISMA 223
 QY 361 TPVAVSGVALLQAHPSWTPDKXTALIEFADIVAPKEIADIYAGAGRVNYKA 434
 Db 224 SEHVAGAAALILSKHPTMTNASQVRNRLSTATYIG---SSFYGKGLINVEAA 273
 RESULT 11
 SUBSI
 subtlisin (EC 3.4.21.62) E precursor - *Bacillus subtilis*
 N/Alternate names: alkaline proteinase; bacillopeptidase E; extracellular alkaline serin
 C/Species: *Bacillus subtilis*
 C/Date: 04-Dec-1986 #sequence, revision 04-Dec-1986 #text_change 05-Oct-2004
 C/Accession: A00972; A26116; I39770; I39778; I39779; S68012; H65586
 R/Stahl, M.L.; Ferrari, E.
 J. Bacteriol. 158, 411-418, 1984

A/Title: Replacement of the *Bacillus subtilis* subtlisin structural gene with an in vitro
 A/Reference number: A00972; MUID:84212198; PMID:6427178
 A/Accession: A00972
 A/Molecule type: DNA
 A/Residues: 1-381 <STA>
 A/Cross-references: UNIPROT:P04189; UNIPARC:UPI0000142525; GB:K01988; NID:g143519; PIDN:
 A/Experimental source: strain 1168
 R/Mong, S.L.; Price, C.W.; Goldfarb, D.S.; Doi, R.H.
 Proc. Natl. Acad. Sci. U.S.A. 81, 1184-1188, 1984
 A/Title: The subtlisin E gene of *Bacillus subtilis* is transcribed from a sigma37 promote
 A/Reference number: A26116; MUID:84144862; PMID:6322190
 A/Accession: A26116
 A/Molecule type: DNA
 A/Residues: 1-155 <MON>
 A/Cross-references: UNIPARC:UPI000016B997; GB:K01443; NID:g143665; PIDN:AAA22814.1; PID:
 R/Ikemura, H.; Takagi, H.; Inouye, M.
 J. Biol. Chem. 262, 7859-7864, 1987
 A/Title: Requirement of pro-sequence for the production of active subtlisin E in *Escheri*
 A/Reference number: I39969; MUID:87222417; PMID:3108260
 A/Accession: I39970
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-156 <IKS>
 A/Cross-references: UNIPARC:UPI000016B956; GB:M16639; NID:g143521; PIDN:AAA22744.1; PID:
 R/Henner, D.J.; Ferrari, E.; Perego, M.; Hoch, J.A.
 J. Bacteriol. 170, 296-300, 1988
 A/Title: Location of the targets of the hpx-97, sacQ32(Hy), and sacQ36(Hy) mutations in
 A/Reference number: I39778; MUID:88086885; PMID:2447063
 A/Accession: I39778
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-8 <HEN>
 A/Cross-references: UNIPARC:UPI000016B7CE; GB:M19125; NID:g142527; PIDN:AAA2245.1; PID:
 R/Park, S.
 J. Bacteriol. 171, 2637-2665, 1989
 A/Title: *Bacillus subtilis* subtlisin gene (apre) is expressed from a sigma-A (sigma-43)
 A/Reference number: I39779; MUID:89213955; PMID:2496113
 A/Accession: I39779
 A/Molecule type: DNA
 A/Residues: 1-13 <PAR>
 A/Cross-references: UNIPARC:UPI000016B7CF; GB:M31060; NID:g142529; PIDN:AAA2246.1; PID:
 A/Experimental source: strain W168, substrain PY79
 R/Kamal, M.; Hoeseg, J.O.; Kaiser, R.; Shafiqat, J.; Razzaki, T.; Zaidi, Z.H.; Joernvall,
 FEBS Lett. 374, 363-366, 1995
 A/Title: Isolation, characterization and structure of subtlisin from a thermoestable *Bac*
 A/Reference number: S68012; MUID:96069945; PMID:7589571
 A/Accession: S68012
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 113-323 <KAM>
 A/Cross-references: UNIPARC:UPI0000172C13
 R/Kuner, F.; Ogasawara, N.; Mosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertier
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi
 A.; Enllich, S.D.; Emerson, P.T.; Eutelan, K.D.; Birlington, J.; Febvre, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gallier
 Iech, J.; Harwood, C.R.; Hensat, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;
 Koester, P.; Koningstein, S.; Krogh, S.; Kumano, M.; Kurita, K.; Lepidus, A.; Lardinois,
 A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Meude, S.; Mausel,
 Y.; M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
 A/Authors: Schlögl, S.; Schöfner, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Seter,
 akuchl, M.; Tamakoshi, A.; Tanaka, T.; Terasera, P.; Tognoni, A.; Tozato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K.
 A/Authors: Yoshikawa, H.F.; Zumeireh, E.; Yoshikawa, H.; Danchin, A.
 A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A/Reference number: A69580; MUID:98044033; PMID:9384377
 A/Accession: H69586
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-26, 'V', 28-381 <KUN>
 A/Cross-references: UNIPARC:UPI00006019A; GB:Z99109; GB:AL009126; NID:g2633260; PIDN:CAI
 A/Experimental source: strain 168

C:Comment: Secretion of subtilisin is associated with the onset of sporulation, and many not necessary for normal sporulation.

C:Genetics:

A:Gene: aprE

A:Map position: 690-771

A:Start codon: GTG

C:Function:

A:Description: catalyzes the hydrolysis of peptide bonds

A:Pathway: protein digestion

A:Note: this enzyme has broad specificity and will hydrolyze peptide amides; it prefers

C:Superfamily: Subtilisin; subtilisin homology

C:Keywords: extracellular protein; hydrolase; protein digestion; serine proteinase

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-106/Domain: activation peptide #status predicted <APT>

F:107-381/Product: subtilisin E #status predicted <MAT>

F:129-341/Domain: subtilisin homology <SBT>

F:138,170,327/Active site: Asp, His, Ser #status predicted

Query Match 15.3%; Score 523; DB 1; Length 381;

Best Local Similarity 37.6%; Pred. No. 1.5e-21;

Matches 144; Conservative 54; Mismatches 127; Indels 58; Gaps 15;

```

QY 61 TVIVENHREKEIAVRVLELMGAKRYVYHIIIPALADL--KRDLLVLSGLTGKATL 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 46 TWSAMSSAKKCD---VISEKGKVKQKPKYVNAATAATDEKAVKEL-----KRD 91
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 SGVRFIOEDYKVTSAE-LEGDESAQVMATYVNNLGYDGSIGITIGITIGIDASHPTL 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 92 PSVAAYVEEDH---IAHEVAQSVPYGISQIKAPALHSQGTGSNVKVAVIDSGIDSHPTL 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 177 QSKVIGWDFVNGRSYPYDD-HGHGTHVASIAAGTGAANGK--YKGMAPGAKLAGIKYL 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 149 --NVRGASFPVSENPYPQDSSHGTHV---AGTIALNNSIGVLGVSPASLVAVKYL 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 234 GADGSGSISTIIKGYEAVDNKDKYGIKVINLSLSSGSGSDGTDLSQAVNNAMPAGIYV 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 203 DSTGSGQYSWIINGIEMAISN---NMDVINNSLIG---GPTGSTALKTVVDAVSSGIYV 255
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 294 CVAANGSGP--NTYVGSPPAAASKVITVGAVDSNDNIASFSSRGPTADRLKPEVAPGV 351
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 256 AAAAGNESSGSSSTVGIPAKYPTTAVGAVNNSQRPASFSSGSELD-----VWAPGV 309
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 352 DIIAPRAGTSMGTPINDYTYKASGTMATPHVSGVALLIOAHPSWTPDKYKTALIIETA 411
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 310 SIQSTLPQGT-----YGAVNGTSMATPHVAGAAALLILSKHPTWNAQVRDLLESTA 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 412 DIVAPKEIADIYAGGRVNVYKA 434
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 361 TYLG-----NSFYGKGLINVQAA 379
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 12

SUBS subtilisin (BC 3.4.21.62) amylosacchariticus precursor - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 24-Apr-1984 #sequence revision 24-Feb-1995 #text_change 05-Oct-2004

C:Accession: A41448; A00971; S68013

R.Yoshimoto, T.; Oyama, H.; Honda, T.; Tone, H.; Takeshita, T.; Kamiyama, T.; Tsuru, D.

J. Biochem. 103, 1060-1065, 1988

A:Title: Cloning and expression of subtilisin amylosacchariticus gene.

A:Reference number: A41448; MUID:89008194; PMID:3139650

A:Accession: A41448

A:Molecule type: DNA

A:Residues: 1-381 <YOS>

A:Cross-References: UNIPROT:P00783; UNIPARC:UPI0000136184; GB:D00264; NID:g216328; PIND:

A:Experimental source: var. amylosacchariticus

R.Kurihara, M.; Markland, F.S.; Smith, E.L.

J. Biol. Chem. 247, 5619-5631, 1972

A:Title: Subtilisin amylosacchariticus. III. Isolation and sequence of the chymotryptic

A:Reference number: A00971; MUID:7226688; PMID:5055784

A:Accession: A00971

A:Molecule type: protein

A:Residues: 107,112-114;148-152;155-157,164-170;173-174,178-181;200-205;210-212;219-225;

A:Cross-References: UNIPARC:UPI000011894; UNIPARC:UPI00001708CD; UNIPARC:UPI0000172B66; C07; UNIPARC:UPI0000172C08; UNIPARC:UPI0000172C09; UNIPARC:UPI0000172C0A; UNIPARC:UPI0000172C10; UNIPARC:UPI0000172C11

A:Experimental source: var. amylosacchariticus

R.Kamel, M.; Hoeoeg, J.O.; Kaiser, R.; Shafiq, J.; Razaki, T.; Zaidi, Z.H.; Joernvall,

FEBS Lett. 374, 363-366, 1995

A:Title: Isolation, characterization and structure of subtilisin from a thermostable Bac

A:Reference number: S68012; MUID:96069945; PMID:7589571

A:Accession: S68013

A:Status: preliminary

A:Molecule type: protein

A:Residues: 107-235, 'T', 237-245, 293-381 <KAM>

A:Cross-References: UNIPARC:UPI0000172C12

C:Comment: Secretion of subtilisin is associated with the onset of sporulation, and many

not necessary for normal sporulation.

C:Genetics:

A:Start codon: GTG

C:Superfamily: Subtilisin; subtilisin homology

C:Keywords: extracellular protein; hydrolase; serine proteinase

F:107-381/Product: subtilisin #status predicted <MAT>

F:129-341/Domain: subtilisin homology <SBT>

F:138,170,327/Active site: Asp, His, Ser #status predicted

Query Match 15.2%; Score 520; DB 1; Length 381;

Best Local Similarity 37.3%; Pred. No. 2.3e-21;

Matches 143; Conservative 55; Mismatches 127; Indels 58; Gaps 15;

```

QY 61 TVIVENHREKEIAVRVLELMGAKRYVYHIIIPALADL--KRDLLVLSGLTGKATL 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 46 TWSAMSSAKKCD---VISEKGKVKQKPKYVNAATAATDEKAVKEL-----KRD 91
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 SGVRFIOEDYKVTSAE-LEGDESAQVMATYVNNLGYDGSIGITIGITIGIDASHPTL 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 92 PSVAAYVEEDH---IAHEVAQSVPYGISQIKAPALHSQGTGSNVKVAVIDSGIDSHPTL 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 177 QSKVIGWDFVNGRSYPYDD-HGHGTHVASIAAGTGAANGK--YKGMAPGAKLAGIKYL 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 149 --NVRGASFPVSENPYPQDSSHGTHV---AGTIALNNSIGVLGVSPASLVAVKYL 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 234 GADGSGSISTIIKGYEAVDNKDKYGIKVINLSLSSGSGSDGTDLSQAVNNAMPAGIYV 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 203 DSTGSGQYSWIINGIEMAISN---NMDVINNSLIG---GPTGSTALKTVVDAVSSGIYV 255
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 294 CVAANGSGP--NTYVGSPPAAASKVITVGAVDSNDNIASFSSRGPTADRLKPEVAPGV 351
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 256 AAAAGNESSGSSSTVGIPAKYPTTAVGAVNNSQRPASFSSGSELD-----VWAPGV 309
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 352 DIIAPRAGTSMGTPINDYTYKASGTMATPHVSGVALLIOAHPSWTPDKYKTALIIETA 411
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 310 SIQSTLPQGT-----YGAVNGTSMATPHVAGAAALLILSKHPTWNAQVRDLLESTA 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 412 DIVAPKEIADIYAGGRVNVYKA 434
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 361 TYLG-----NSFYGKGLINVQAA 379
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 13

subtilisin (BC 3.4.21.62) J precursor - Bacillus stearothermophilus

C:Species: Bacillus stearothermophilus

C>Date: 17-Jul-1992 #sequence revision 17-Jul-1992 #text_change 05-Oct-2004

C:Accession: J01487

R.Jiang, J.S.; Kang, D.O.; Chun, M.U.; Byun, S.M.

Biochem. Biophys. Res. Commun. 184, 277-282, 1992

A:Title: Molecular cloning of a subtilisin J gene from Bacillus stearothermophilus and it

A:Reference number: J01487; MUID:92231938; PMID:1567435

A:Accession: J01487

A:Molecule type: DNA

A:Residues: 1-381 <JAN>

A:Cross-References: UNIPROT:P29142; UNIPARC:UPI0000136185; GB:M64743; NID:g142531; PIND:

A:Experimental source: strain NCIMB10278

C:Genetics:

A:Gene: aprU

A:Start codon: GTG
C:Superfamily: Subtilisin; subtilisin homology
C:Keywords: hydrolase; serine proteinase; zymogen
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-106/Domain: activation peptide #status predicted <ACT>
F:107-381/Product: subtilisin 3 #status predicted <MAT>
F:129-341/Domain: subtilisin homology <SBT>
F:138-170/Active site: Asp, His, Ser #status predicted

Query Match 15.2%; Score 520; DB 2; Length 381;
Best Local Similarity 37.3%; Pred. No. 2,3e-21;
Matches 143; Conservative 55; Mismatches 127; Indels 58; Gaps 15;

QY 61 TVIVFENHREKEIAVRVLEMGAKRVYVYHIIPALADL--KYRDLVLSGLTGKAKL 117
DB 46 TWGAMSAAKKKD-----VISEKGVQKQFVYVNAATAATLEKAVKEL-----KKD 91
QY 118 SGVRFIOEDYKVTVAE-LEGDESAQAQVATYVNLGYDGGITIGITDGDASHPLD 176
DB 92 PSVAAYVEEDH---IAHFAQSVPYGISQIRAPALHSQYGSVNVKVAVIDSGIDSSHPDL 148
QY 177 QGVVIGVNDVFNRSYVYD-HHGHGVASIAAGTGAASNGK--YKGMARGAKLGIKVL 233
DB 149 -NVRGASFPSEETNYQDSSSHGHV---AGTIALNNSIGVLGVSPASLVAVKVL 202
QY 234 GADSGSISTIIKGVEMAVNDKDKYIKVNLSTGSSQSDGTDSLQAANNMADAGIVV 233
DB 203 DSGSGQGYWINGIEIAIN---NMDVIMSLG---GPSGSFALKTVVDKAVSSGIVV 255
QY 294 CVAAGNSGP--NTYTGSPAAASKVITVGAVDSDNDNIASFSSRGPADGRLKEVVAAPGV 351
DB 256 AAAAGNNGSSGSSSTVGYPAKYPSTTIAVGAVNSNMQPASFSSAGSELD-----VVAAPGV 309
QY 352 DIIAPRASGTSWGPINDYTTKASGTSMATPHVSGVALLQAHSPMTPOKVTALILETA 411
DB 310 SIQSTLPFGH-----YGAVNGISMATPHVGAALILSKHPVTMAQVDRLESTA 360
QY 412 DIVAPKEIADIAYGAGRVNYKA 434
DB 361 TYLG---NSFYGKGLINVOAA 379

RESULT 14
A48373
high-alkaline serine proteinase (EC 3.4.21.-) precursor - *Bacillus* sp. (strain AH-101)
N/Alternate names: subtilisin-like thermostable alkaline serine proteinase
C/Species: *Bacillus* sp.
C/Date: 03-Feb-1994 #sequence revision 03-Feb-1994 #text change 05-Oct-2004
C/Accession: A48373; J050714
R/Takami, H.; Kobayashi, T.; Aono, R.; Horikoshi, K.
Appl. Microbiol. Biotechnol. 38, 101-108, 1992
A/Title: Molecular cloning, nucleotide sequence and expression of the structural gene for subtilisin from *Bacillus* sp. strain AH-101
A/Reference number: A48373; MUID:93089926; PMID:11369007
A/Accession: A48373
A/Molecule type: DNA
A/Residues: 1-361 <TA>
A/Cross-references: UNIPARC:UPI0000175C86; GB:SS0880; NID:g2e1737; PIDM:AA060421.1; PID:
A/Experimental source: AH-101
A/Note: this sequence is inconsistent with the nucleotide translation
R/Takami, H.; Kobayashi, T.; Aono, R.; Horikoshi, K.
submitted to JIPID, July 1992
A/Description: Molecular cloning, nucleotide sequence and expression of the structural gene for subtilisin from *Bacillus* sp. strain AH-101
A/Reference number: J050714
A/Accession: J050714
A/Molecule type: DNA
A/Residues: 94-334, 'U', 336-361 <TA2>
A/Cross-references: UNIPARC:UPI0000175C87
C/Comment: This alkaliphilic *Bacillus* homology to the subtilisins of neutrophilic *Bacillus*
C/Superfamily: Subtilisin; subtilisin homology
C/Keywords: extracellular protein; hydrolase; serine proteinase
F:115-321/Domain: subtilisin homology <SBT>
F:124,154,307/Active site: Asp, His, Ser #status predicted

Query Match 14.8%; Score 508; DB 2; Length 361;
Best Local Similarity 33.7%; Pred. No. 9.5e-21;
Matches 147; Conservative 79; Mismatches 128; Indels 82; Gaps 17;

QY 2 KGIKALIVLVGLVGVSAAPAEKQVQVNVNKGVLTPGLRKIQKLNPNNEIST 61
DB 3 QSLKAVLVSTVAL-LPMANPAAASEKRYLVVEP-----DEVSA 42
QY 62 VIIVFENHREKEIAVRVLEMGAKRVYVYHIIPALADLKYRDLVLSGLTGKAKLSGV 121
DB 43 QSVSESY-----DVDVHFEFEPVPIHLELBEELKELQNDPNVYA----- 83
QY 122 FIOEDYKVTVAE-LEGDESAQAQVATYVNLGYDGGITIGITDGDASHPLDQGV 180
DB 84 -IENAEVITISQVPMGISFISTQ---QAHNRGIFGNARVAVLDTGI-ASHPLD--R1 135
QY 181 IGVNDVFNRSYVYDHHGHGVASIAAGTGAASNGK--YKGMARPAKLGIKVLGADSD 238
DB 136 AGGASFPSEPSYHDNNGHGHV---AGTIALNNSIGVLGVAPADLVAVKVLDRNGS 191
QY 239 GSISTIIKGVEMAVNDKDKYIKVNLSTGSSQSDGTDSLQAANNMADAGIVVCAAG 298
DB 192 GSLSAVAQIGIEMALIN---NHAIIMSLG---STGSSFTELVAVRANAGILLVGAAG 244
QY 299 NSGPNYTVGSPAAASKVITVGAVDSDNDNIASFSSRGPADGRLKEVVAAPGVDTIAPRA 358
DB 245 NTG--RQGVNYPARYGVMAVAADQNGQASFSYGP-----EIEIAPGVNVYS-RY 295
QY 359 SGTSMGTPINDYTTKASGTSMATPHVSGVALLQAHSPMTPOKVTALILETDIAPKE 418
DB 296 TG-----NRRVLSGTSMATPHVGAALVKSRYSPTNNQIRINQITATYLGSSN 347
QY 419 IADIAYGAGRVNYKA 434
DB 348 L-----YGNGLVHAGRA 359

RESULT 15
G83756
subtilisin-type alkaline proteinase (EC 3.4.21.-) BH0855 precursor [similarity] - *Bacilli*
C/Species: *Bacillus* halodurans
C/Date: 01-Dec-2000 #sequence revision 01-Dec-2000 #text change 05-Oct-2004
C/Accession: G83756
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiram
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus* halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: G83756
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-361 <STO>
A/Cross-references: UNIPARC:PA1363; UNIPARC:UPI0000129E77; GB:AP001510; GB:BA000004; NID:
A/Experimental source: strain C-125
A/Genes: BH0855
C/Superfamily: Subtilisin; subtilisin homology
C/Keywords: hydrolase; serine proteinase
F:1-25/Domain: signal sequence #status predicted <SIG>

Query Match 14.8%; Score 508; DB 2; Length 361;
Best Local Similarity 33.1%; Pred. No. 9.5e-21;
Matches 144; Conservative 80; Mismatches 131; Indels 80; Gaps 15;

QY 2 KGIKALIVLVGLVGVSAAPAEKQVQVNVNKGVLTPGLRKIQKLNPNNEIST 61
DB 3 QSLKAVLVSTVAL-LPMANPAAASEKRYLVVEP-----DEVSA 42
QY 62 VIIVFENHREKEIAVRVLEMGAKRVYVYHIIPALADLKYRDLVLSGLTGKAKLSGV 121
DB 43 QSVSESY-----DVDVHFEFEPVPIHLELBEELKELKUKL-----KDPNVK 82
QY 122 FIOEDYKVTVAE-LEGDESAQAQVATYVNLGYDGGITIGITDGDASHPLDQGV 181

```

Db      83 ALEKXAEVTIS---QTVPMGISFINTQQAHHNRGIFGNGARVAVLDTGI-ASHPDL--RIA 136
QY      182 GWVDEVNGRSYPPYDDHGHGTHVASTAAGTGAASNGK--YKGMAPGAKLAGIKVLGADGSG 239
Db      137 GGASPISEPSYHDNNGHGHV---AGTIALNNISIGVLGVAFPSADLYAVKVLDRNGSG 192
QY      240 SISTIKGVEMAVDNKDKYGIKVINLSIGSSQSSDQDTSLSQAVNNAMADAGIVCVAAGN 299
Db      193 SIASVAQGIEMAHNN---NMHIIIMSIG--STGSSSTLELAANRANAGILIVGAAGN 245
QY      300 SGPNTYTVGSPAAASKVITVGAVDSNDNIASFSSRGPTADGRLKPEVVA PGVDIIAPRAS 359
Db      246 TG--RQGVNYPARYSGVMAVAAPDQNGORASFSTYGP-----EIEISAPGVN----- 291
QY      360 GTSMTGPIINDYITKASGTSMTAPHVSGVICALILOAHPSTPDKXTALLETADIAPKEI 419
Db      292 ---NSTYTGNRVSLSGTSMATPHVAGVAAALVKSRYPSTYNNQIRORINOTATYLGSPSL 348
QY      420 ADIAYGAGRNVVYKA 434
Db      349 ----YNGGLVHAGRA 359

```

Search completed: January 6, 2007, 22:32:20
 Job time : 29 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2007 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2007, 22:20:17 ; Search time 192 Seconds

(without alignments)
3174.923 Million cell updates/sec

Title: US-10-800-684-5
Perfect score: 3428
Sequence: 1 MKGLKALIVILVGLVVGVS.....YAYSTYGMADYQLKAVYVYG 659

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : UniProt 7.2.*
1: uniprot_sprot.*
2: uniprot_trcml.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------|
| 1 | 3015.5 | 88.0 | 654 | 2 | 08U0C9_PYRFU |
| 2 | 2705 | 78.9 | 663 | 2 | 05J1Z5_PYRKO |
| 3 | 980.5 | 28.6 | 561 | 2 | 08RB02_THETN |
| 4 | 925 | 27.0 | 424 | 2 | 03C085_THESH |
| 5 | 752 | 21.9 | 1294 | 2 | 050HM7_STRSH |
| 6 | 748.5 | 21.8 | 1245 | 2 | 05RL54_STRCO |
| 7 | 737.5 | 21.5 | 795 | 2 | 05NM24_9ARCH |
| 8 | 715.5 | 20.9 | 430 | 2 | 08ENV1_OCBTH |
| 9 | 703 | 20.5 | 1139 | 2 | 082139_STRAV |
| 10 | 702 | 20.5 | 1105 | 2 | 08KXh6_STRVD |
| 11 | 688 | 20.1 | 442 | 2 | 065IP4_BACSD |
| 12 | 688 | 20.1 | 442 | 2 | 031788_BACSD |
| 13 | 681.5 | 19.9 | 442 | 2 | 05L315_GEOKA |
| 14 | 674.5 | 19.7 | 412 | 2 | 03CGT3_THETH |
| 15 | 668.5 | 19.5 | 1102 | 2 | 035684_STRAO |
| 16 | 667.5 | 19.5 | 444 | 2 | 09KXJ7_BACHD |
| 17 | 664.5 | 19.4 | 1220 | 2 | 09KXJ7_STRCO |
| 18 | 662.5 | 19.3 | 1208 | 2 | 082B14_STRAM |
| 19 | 658 | 19.2 | 1253 | 2 | 09FC06_STRCO |
| 20 | 655 | 19.1 | 1237 | 2 | 08GCT4_STRAZ |
| 21 | 644.5 | 18.8 | 412 | 2 | 08R668_THETN |
| 22 | 642.5 | 18.7 | 412 | 2 | 09AER6_THETO |
| 23 | 635.5 | 18.5 | 369 | 2 | 03C523_9CIOT |
| 24 | 635.5 | 18.5 | 644 | 2 | 04G621_METBA |
| 25 | 634.5 | 18.5 | 435 | 2 | 08EMJ3_OCEIH |
| 26 | 615.5 | 18.0 | 1239 | 2 | 08PM24_STRCO |
| 27 | 564 | 16.5 | 374 | 2 | 09F941_BACLI |
| 28 | 562 | 16.4 | 379 | 2 | 06BCN9_BACKO |
| 29 | 562 | 16.4 | 379 | 2 | 09FDF4_BACLI |
| 30 | 562 | 16.4 | 379 | 2 | 06SLP7_BACLI |
| 31 | 561 | 16.4 | 374 | 2 | 09F942_BACLI |

| | | | | | | |
|----|-------|------|------|---|--------------|--------------------|
| 32 | 560 | 16.3 | 379 | 2 | 06PNN5_BACLI | 06PNN5 bacillus 11 |
| 33 | 559 | 16.3 | 374 | 2 | 09F943_BACLI | 09F943 bacillus 11 |
| 34 | 559 | 16.3 | 379 | 2 | 04PKR6_BACLI | 04PKR6 bacillus 11 |
| 35 | 556 | 16.2 | 379 | 2 | 053521_BACLI | 053521 bacillus 11 |
| 36 | 555 | 16.2 | 382 | 1 | SUBT_BACAM | P00782 bacillus am |
| 37 | 554 | 16.2 | 379 | 1 | SUBT_BACLI | P00780 bacillus 11 |
| 38 | 551 | 16.1 | 379 | 2 | 045300_BACLI | 045300 bacillus 11 |
| 39 | 546.5 | 15.9 | 1358 | 2 | 08ETM4_OCEIH | 08ETM4 oceanobacil |
| 40 | 545 | 15.9 | 379 | 2 | 045301_BACLI | 045301 bacillus 11 |
| 41 | 544 | 15.9 | 382 | 2 | 045522_9BACI | 045522 bacillus sp |
| 42 | 543.5 | 15.9 | 310 | 2 | 09FDF3_BACLI | 09FDF3 bacillus 11 |
| 43 | 543.5 | 15.9 | 1398 | 2 | 09F911_PYRMO | 09F911 pyrococcus |
| 44 | 542.5 | 15.8 | 567 | 2 | 048811_COLP3 | 048811 colwellia p |
| 45 | 541.5 | 15.8 | 1398 | 1 | PLS_PYRFU | P72186 pyrococcus |

ALIGNMENTS

| RESULT 1 | 08U0C9_PYRFU | PRELIMINARY; | PRT; | 654 AA. |
|-----------------------|---|--------------|------|---------|
| AC | 08U0C9; | | | |
| DT | 01-JUN-2002, integrated into UniProtKB/TrEMBL. | | | |
| DT | 01-JUN-2002, sequence version 1. | | | |
| DT | 07-FEB-2006, entry version 19. | | | |
| DE | Alkaline serine protease. | | | |
| GN | OrderedAccessionName=PI670; | | | |
| OS | Pyrococcus furiosus. | | | |
| OC | Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae; | | | |
| OC | Pyrococcus. | | | |
| OX | NCBI_TaxID=2261; | | | |
| RN | [1] | | | |
| RP | NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | | | |
| RC | STRAIN=Vci / DSM 3638 / ATCC 43587 / JCM 8422. | | | |
| RA | Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.; | | | |
| RT | "The complete sequence of the Pyrococcus furiosus genome." | | | |
| RL | Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases. | | | |
| CC | Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms | | | |
| CC | Distributed under the Creative Commons Attribution-NoDerivs license | | | |
| DR | EMBL; AE010265; AAL81794.1; -; Genomic DNA. | | | |
| DR | HSSP; Q99405; 1MPT. | | | |
| DR | BioCyc; PPR186497:PI670-MONOMER; -. | | | |
| DR | GO; GO:0008233; F:peptidase activity; IEA. | | | |
| DR | GO; GO:0004289; F:subtilase activity; IEA. | | | |
| DR | GO; GO:0006508; P:proteolysis; IEA. | | | |
| DR | InterPro; IPR007280; Pept arc bac_C. | | | |
| DR | InterPro; IPR000209; Pept_S8_S53. | | | |
| DR | Pfam; PF00082; Peptidase_S8; 1. | | | |
| DR | Pfam; PF04151; PFC; 2. | | | |
| DR | PRINTS; PR00723; SUBTILASIN. | | | |
| DR | PROSITE; PS00136; SUBTILASE ASP; UNKNOWN_1. | | | |
| DR | PROSITE; PS00137; SUBTILASE HIS; UNKNOWN_1. | | | |
| DR | PROSITE; PS00138; SUBTILASE_SER; 1. | | | |
| KW | Complete proteome. | | | |
| SQ | SEQUENCE 654 AA; 70231 MW; 1CB145A5FE50DB34 CRC64; | | | |
| Query Match | 88.0%; Score 3015.5; DB 2; Length 654; | | | |
| Best Local Similarity | 88.8%; Pred. No. 1.9e-157; | | | |
| Matches | 585; Conservative 22; Mismatches 47; Indels 5; Gaps 1; | | | |
| QY | 1 MKGLKALIVILVGLVVGSAAPDKYQVRYNKGKYLTPGLFKRIQKLNPEEIS 60 | | | |
| DB | 1 MKGLKALIVILVGLVVGSAAPDKYQVRYNKGKYLTPGLFKRIQKLNPEEIS 60 | | | |
| QY | 61 TYIVENHREKEIAVRVLEMGAKVRYVYHIIIPALADIKVLDLVISGLTGSKAKLSGV 120 | | | |
| DB | 61 TYIVENHREKEIAVRVLEMGAKVRYVYHIIIPALADIKVLDLVISGLTGSKAKLSGV 120 | | | |
| QY | 121 REIQDYKVTYSABLEGIDESAAYQWATYVNNLGYDGGITTIIGIDTGIDASHPLQSKV 180 | | | |

Db 121 RFIQEDYKTVASAELEGLDESAQWMAITYWNILGYDSSGTTIGIITDGDASHEDLQKGV 180
 QY 181 IGWVDFVNGRSYPYDDHGHGTHVASIAGTGAASNGKYKGMAPGAKIAGIKVLGAADSSG 240
 Db 181 IGWVDFVNGRSYPYDDHGHGTHVASIAGTGAASNGKYKGMAPGAKIAGIKVLGAADSSG 240
 QY 241 ISTTIKVENAVNDKDKYKIGKIVNLISGSSQSSDGTDSLQAANNAMDAQIIVCVAAAGNS 300
 Db 241 ISTTIKVENAVNDKDKYKIGKIVNLISGSSQSSDGTDSLQAANNAMDAQIIVCVAAAGNS 300
 QY 301 GPNVTYVGSPPAAASKVITTVGAVDNDNIASFSSRGPTADRLKPEVVAPEGVDIIAPASG 360
 Db 301 GPNVTYVGSPPAAASKVITTVGAVDNDNIASFSSRGPTADRLKPEVVAPEGVDIIAPASG 360
 QY 361 TSMGQPIINDYTTAAGTSMATPHVAGIALLLQAHPSMTDPDKYKIALIETADIYKDEIA 420
 Db 361 TSMGQPIINDYTTAAGTSMATPHVAGIALLLQAHPSMTDPDKYKIALIETADIYKDEIA 420
 QY 421 DIAYGAGRVVYVYAIKTDYAKLTFTGSVADKGSATHTPVSQATFTATLVYNDTGSSTI 480
 Db 421 DIAYGAGRVVYVYAIKTDYAKLTFTGSVADKGSATHTPVSQATFTATLVYNDTGSSTI 480
 QY 481 DLVLDPNGNEVYVYATYAGFEKVGYYNPAGTWTVKVSYKGAANYQVDVYDGLSLQ 540
 Db 481 DLVLDPNGNEVYVYATYAGFEKVGYYNPAGTWTVKVSYKGAANYQVDVYDGLSLQ 540
 QY 541 SGGPNPNPNPNPTPTTDTQFTGTVNDYDTSDFTMNVNSGATKITGDLTFDTSYND 600
 Db 541 SGGPNPNPNPNPTPTTDTQFTGTVNDYDTSDFTMNVNSGATKITGDLTFDTSYND 600
 QY 601 LDLYLDPNGNVLVDRSTSSNSYEHVEYANPAGTWTFLVYASTYGMADYQLKAVYVYG 659
 Db 601 LDLYLDPNGNVLVDRSTSSNSYEHVEYANPAGTWTFLVYASTYGMADYQLKAVYVYG 659
 QY 596 LDLYLDPNGNVLVDRSTSSNSYEHVEYANPAGTWTFLVYASTYGMADYQLKAVYVYG 654
 Db 596 LDLYLDPNGNVLVDRSTSSNSYEHVEYANPAGTWTFLVYASTYGMADYQLKAVYVYG 654
 RESULT 2
 OSJ125_PYRKO PRELIMINARY; PRT; 663 AA.
 AC OSJ125;
 DT 15-FEB-2005, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 1.
 DE Subtilisin-like serine protease.
 GN OrderedlocusNames=TK1689;
 OS Pyrococcus kodakaraensis (Thermococcus kodakaraensis).
 OC Archaea; Euryarchaeota; Thermococci; Thermococcaceae;
 OC Thermococcus.
 NCBI_TaxID=69014;
 RN NCBI_TaxID=69014;
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=KOD1;
 RA Fukui T., Atomi H., Kanai T., Matsumi R., Fujiwara S., Imanaka T.;
 RX PubMed=15710748; DOI=10.1101/gr.3003105;
 RT "Complete genome sequence of the hyperthermophilic archaeon
 Thermococcus kodakaraensis KOD1 and comparison with Pyrococcus
 genomes";
 RL genome Res. 15:352-363 (2005).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonCommercial License
 CC EMBL; AP006878; BAD85878.1; -, Genomic DNA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilisin-like serine protease activity; IEA.
 DR GO; GO:0006508; P:proteolysis; IEA.
 DR InterPro; IPR007280; Pept_arc_bac_C.
 DR InterPro; IPR000209; Pept_S8_S53.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04511; PFC; 2.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; UNKNOWN 1.
 DR PROSITE; PS00137; SUBTILASE HIS; UNKNOWN 1.
 DR PROSITE; PS00138; SUBTILASE_SRR; 1.

KM Complete proteome; Protease.
 SQ SEQUENCE 663 AA; 70955 MW; 2CE68ACD3888390E CRC64;
 Query Match 78.9%; Score 2705; DB 2; Length 663;
 Best Local Similarity 77.7%; Pred.No.2,3e-140;
 Matches 517; Conservative 60; Mismatches 80; Indels 8; Gaps 4;
 QY 1 MKGLKALIVLIVLGLVGSVAAAPKKVEQVYANV--EKNGYGLTGTGLPRKIOKLPNEE 58
 Db 1 MKKFGAVLVLFLVGLMASSVLAAPKPA--VRNVSOQKNYGLTGTGLRKQRMWDDE 58
 QY 59 ISTVIFENHREKEIAVRLVLEMGAKRVYVHTIIPALADLKVRDLVTSGL--TG--GK 114
 Db 59 VSTTIWFEDNQAEKAEVILDFGAKIKYNYHIIPALAVIKIKODLIIAGLMDTGYFGN 118
 QY 115 AKLSGRFIOEDYKTVASAELEGLDESAQWMAITYWNILGYDSSGTTIGIITDGDASHP 174
 Db 115 AQLSGVQFIQEDYVAVAEETEGLEDSAAQWMAITNNILGYDSSGTTIGIITDGDASHP 178
 QY 175 DLQKVIIGWVDFVNGRSYPYDDHGHGTHVASIAGTGAASNGKYKGMAPGAKIAGIKVLG 234
 Db 175 DLQKVIIGWVDFVNGRSYPYDDHGHGTHVASIAGTGAASNGKYKGMAPGAKIAGIKVLN 238
 QY 235 ADGSGSISTTIKVENAVNDKDKYKIGKIVNLISGSSQSSDGTDSLQAANNAMDAQIIVC 294
 Db 235 GQSGSISDIIINGVMAVQNDKDKYKIGKIVNLISGSSQSSDGTDSLQAANNAMDAQIIVV 298
 QY 295 VAAAGNPNPNPNPNPTPTTDTQFTGTVNDYDTSDFTMNVNSGATKITGDLTFDTSYND 354
 Db 295 VAAAGNPNPNPNPNPTPTTDTQFTGTVNDYDTSDFTMNVNSGATKITGDLTFDTSYND 358
 QY 355 APASGTSMTPIINDYTTAAGTSMATPHVAGIALLLQAHPSMTDPDKYKIALIETADIY 414
 Db 355 APASGTSMTPIINDYTTAAGTSMATPHVAGIALLLQAHPSMTDPDKYKIALIETADIY 418
 QY 415 ARKEIADIYAGRVVYVYAIKTDYAKLTFTGSVADKGSATHTPVSQATFTATLVYNDTGSSTI 474
 Db 415 KPEIADIYAGRVVYVYAIKTDYAKLTFTGSVADKGSATHTPVSQATFTATLVYNDTGSSTI 478
 QY 475 TQSSDIDLVLDPNGNEVYVYATYAGFEKVGYYNPAGTWTVKVSYKGAANYQVDVYDGLSLQ 534
 Db 475 NGSDIDLVLDPNGNEVYVYATYAGFEKVGYYNPAGTWTVKVSYKGAANYQVDVYDGLSLQ 538
 QY 535 DGLSLQSGGPNPNPNPNPTPTTDTQFTGTVNDYDTSDFTMNVNSGATKITGDLTFDTSYND 594
 Db 535 DGLSLQSGGPNPNPNPNPTPTTDTQFTGTVNDYDTSDFTMNVNSGATKITGDLTFDTSYND 598
 QY 595 DTSYNDLVLDPNGNVLVDRSTSSNSYEHVEYANPAGTWTFLVYASTYGMADYQLKAVYVYG 654
 Db 595 DTSYNDLVLDPNGNVLVDRSTSSNSYEHVEYANPAGTWTFLVYASTYGMADYQLKAVYVYG 658
 QY 655 VVYVYG 659
 Db 655 KYYVYG 663
 RESULT 3
 OSRBJ2_THETN PRELIMINARY; PRT; 561 AA.
 AC OSRBJ2;
 DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 1.
 DE Subtilisin-like serine protease.
 GN Name=APRE2; OrderedlocusNames=TTB0824;
 OS Thermomicrobacter tengcongensis.
 OC Bacteria; Firmicutes; Clostridia; Thermomicrobacteriales;
 OC Thermomicrobacteriaceae; Thermomicrobacter.
 NCBI_TaxID=119072;
 RN NCBI_TaxID=119072;
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=MB4 / JCM 11007;
 RX MEDLINE=21992816; PubMed=11997336; DOI=10.1101/gr.219302;

BAO Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;
 RT "A complete sequence of the T. tengcongensis genome.";
 RL Genome Res. 12:689-700(2002).

CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC -----
 DR EMBL: AE013049; AAM24081.1; -; Genomic_DNA.
 DR HSP: P00782; 2SRT.
 DR Biocyc: TTN119072:TTB0824-MONOMER; -;
 DR GO: GO:0008233; F:peptidase activity; IEA;
 DR GO: GO:0042802; F:protein self binding; IEA.
 DR GO: GO:004289; F:subtilase activity; IEA.
 DR GO: GO:0043086; F:negative regulation of enzyme activity; IEA.
 DR GO: GO:0006508; F:proteolysis; IEA.
 DR InterPro: IPR002048; EF_hand_Ca_bd.
 DR InterPro: IPR007280; Pept_arc_bac_C.
 DR InterPro: IPR000209; Pept_s8_S53.
 DR InterPro: IPR010259; Prot_inh_S8A.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR Pfam: PF04151; PC; 1.
 DR Pfam: PF05922; Subtilisin_N; 1.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00018; EF_HAND_1; UNKNOWN_1.
 DR PROSITE: PS00136; SUBTILASE_ASP; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 DR Complete proteome; Hydrolyase; Protease; Serine protease.
 KW SEQUENCE 561 AA; 59696 MW; BA9C5C52F7083A18 CRC64;

Query Match 28.6%; Score 980.5; DB 2; Length 561;

Best Local Similarity 42.3%; Pred. No. 1.3e-45; Matches 246; Conservative 83; Mismatches 174; Indels 79; Gaps 18;

2 KGLKLLIIVLVGLVGVAAPEKVEQVRVEKNYGLTPGLPRKI--QKLPN-- 56
 8 KLLSLALISLISLNEIIVQAPNINLPIDSPKITY-----PSLPQKISMDSNKNKIF 63
 57 -----EISTVIYVEENREKEIAVRVLELWGA-KRVYVHIIPAIADLKVR 102
 64 DDLQRLINKPDESEPPVILITFNKIPVSDADIFTIAKNIGKFNKHKYKIIPIAANLTKS 123
 103 DLIVISGLTGAKAKISGARFIQEDRYKYVTSAEI-----GLDESAQVAMTYVMNUGYDS 158
 124 QINVL-----KLEIVQIEYDEYVATLDTATKMFGITRARS-----DFGVTK 168
 159 GTTIGITDGDASHPDQ-GKVIGWVDFVNGSRYPYDGHGHTVASTAAGTGAASNGK 217
 169 NITLIIIDIGIDGNHVLDSGKTIQKMDITNNKTIPTDNGHGHVASTAAGTG-NGNSF 227
 218 YKGMAPGAKIAGIKVLGADSGSISTITIKGVEMAVDNKQYGIKVINLSLGSSOSSDGT 277
 228 YGVAPDALVGIKVLDAANGSGMSVTAGIDMAVQNKOVYGIKVINLSLGSTSDGT 287
 278 SLSQAVNNAMADGIYVCVAANGSGPNTYVGSPPAASKITYTGAV---DSNDINAFSS 333
 288 STSLVNNRAVDGSIYVVVAANGSGPAKYTIGSGAEKAITVAAAMADVLELFPNLSFSS 347
 334 RGPATDGRKLPKPEVVAAGVDIIAPRASGTSMTGPIINDYTKASGTSMAATPHVSGVGLLIQ 393
 348 RGPITDGRKLPKPIAPAGYVITAAKANS-----VNGYVT-YSGTSMATPFVAGTVALMLN 400
 394 AHPSTPDKYKTALETADIAPKEIADIAYGAGRVNYYKAIKYDDYAKLTFTGSVAD-- 451
 401 ANPNILPLNDAKNINISTAKSGPPS-KNVYDYGAGRIDYEAIRVAG-----NFRGNNDIVP 455
 452 -----KGSATHPDVSGARF-VTATLY---WDTGSSDIDLXYDPNGEVNYS 495
 456 NHYISGYLPGSRYSSTJWTFNATNTSYPIATLILIPDMANNPDDFIYLYDSGTLIK-S 514
 496 YTAAYGFEKVGYYNPAGTWTYVVSYKGAANYQVDVSDS 537

DB 515 STGTORQETITLPSQTYVYKVSYSRSGNYPFDLSAGS 556

RESULT 4
 ID Q3CB85 THEBT PRELIMINARY; PRT; 424 AA.
 AC Q3CB85;
 DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
 DT 22-NOV-2005, sequence version 1.
 DT 07-FEB-2006, entry version 4.
 DE Peptidase 58 and 553, subtilisin, kexin, sedolisin, peptidase, archaeal
 DE and bacterial C-terminal.
 GN ORFName=Tech39DRAFT_0431;
 OS Thermoaerobacter ethanolicus ATCC 33223.
 OC Bacteria; Firmicutes; Clostridia; Thermoaerobacteriales;
 OC Thermoaerobacteriaceae; Thermoaerobacter.
 OX NCBI_TaxID=340099;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 33223;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler J.C., Glavina T.,
 RA Hammon N., Israni S., Pittluck S., Richardson P.;
 RT "Sequencing of the draft genome and assembly of Thermoaerobacter
 ethanolicus 39E";
 RL Submitted (SEP-2005) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 33223;
 RG US DOE Joint Genome Institute (JGI-ORNL);
 RA Larimer F., Land M.;
 RT "Annotation of the draft genome of Thermoaerobacter ethanolicus
 39E";
 RL Submitted (OCT-2005) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC -----
 DR EMBL: AAK001000003; BA065215.1; -; Genomic_DNA.
 DR GO: GO:0008233; F:peptidase activity; IEA.
 DR GO: GO:0004283; F:subtilase activity; IEA.
 DR GO: GO:0006508; F:proteolysis; IEA.
 DR Hydrolyase; Protease; Serine protease.
 KW SEQUENCE 424 AA; 44126 MW; 335746D47D09B928 CRC64;

Query Match 27.0%; Score 925; DB 2; Length 424;

Best Local Similarity 51.2%; Pred. No. 1e-42; Matches 210; Conservative 52; Mismatches 112; Indels 36; Gaps 11;

155 YDGSIGITIGIDTGDASHPDQ-GKVIGWVDFVNGSRYPYDGHGHTVASTAAGTGA 213
 24 YEKNDIYAVIDTGDISHVDLAGKRVIGQDFVNGSKSPYDNGHGHVASTAAGTGG 83
 214 SNGKTKMAPGAKIAGIKVLGADSGSISTITIKGVEMAVDNKQYGIKVINLSLGSSOSS 273
 84 NN-LYKGVAPGAALVIGIKVLDNSGSGTMTVAGIDMAVQNDIYGIKVINLSLGSTSS 142
 274 DQDLSLQAVNNAMADGIYVCVAANGSGPNTYVGSPPAASKITYTGAV---DSNDINAF 329
 143 DQDSTSLAVNAGVDSGIVVVVAANGSGPNTYVGSPPAASKITYTGAV---DSNDINAF 202
 330 SPSRSGPTADGRKLPKPEVVAAGVDIIAPRASGTSMTGPIINDYTKASGTSMAATPHVSGVGA 389
 203 SPSRSGPTADGRKLPKPIAPAGYVITAAKANS-----INGYVT-YSGTSMATPFVAGTVA 255
 390 LILQHPSTPDKYKTALETADIAPKEIADIAYGAGRVNYYKAIKYDDYAKLTFTGSVAD-- 439
 256 LMLSANINLAPLDAKNIIMTTAKSGPPS-KNIDYGVGLDAYEAIKTAGNFTGNTISVP 314

```

QY 440 ---YAKITFGSVADGSGATHPEDGGAFFVATLLY----MDTSGSIDLILYDPMGNEV 493
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 NHYAAKSLSGS---RYSDLMTFNVDISYPIAIFILPDMANTPPDDILYDIPGTLLV 371
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 493 DYSTALYGEKGYGNPTAGTWTAVVYSKGAANYQVDV-VSDGSLSGS 541
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 372 K-SITGCRQROETIYLPIDGIIYIKVYSFRSGSNYPDLVSGGSGSLTLS 420
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 5

| | | |
|----|--|---|
| ID | | 05OHM7_STRESH PRELIMINARY; PRt; 1294 AA. |
| AC | | 05OHM7; |
| DT | | 07-JUN-2005, integrated into UniProtKB/TrEMBL. |
| DT | | 07-JUN-2005, sequence version 1. |
| DT | | 07-FEB-2006, entry version 4. |
| D8 | | Probable secreted peptidase. |
| OS | | Streptomycetes sphaeroides. |
| OC | | Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes. |
| OX | | NCB1_taxID=195949; |
| RN | | [1] |
| RP | | NUCLEOTIDE SEQUENCE. |
| RC | | STRAIN=NCIMB 11891; |
| PX | | PubMed=15870333, DOI=10.1128/AEM.71.5.2452-2459.2005; |
| RA | | Eustaquio A.S., Gust B., Galm U., Li S.-W., Chater K.F., Heide L.; |
| RT | | "Heterologous Expression of Novobiocin and Chlorobiotin Biosynthetic Gene Clusters."; |
| RL | | Appl. Environ. Microbiol. 71:2452-2459(2005). |
| RN | | [2] |
| RP | | NUCLEOTIDE SEQUENCE. |
| RC | | STRAIN=NCIMB 11891; |
| RA | | Siefken M., Li S.-W., Heide L.; |
| RL | | Submitted (JAN-2003) to the EMBL/Genbank/DDBJ databases. |
| CC | | Copysighted by the UniProt Consortium, see http://www.uniprot.org/terms Distributed under the Creative Commons Attribution-NoDerivs license |
| CC | | ----- |
| DR | | EMBL; AY227005; AAP4601.1; -; Genomic DNA. |
| GO | | GO:0008233; F:peptidease activity; INA. |
| DR | | GO; GO:0004289; P:protease activity; IEA. |
| GO | | GO:0006508; P:proteinolysis; IEA. |
| DR | | InterPro; IPR003137; PA. |
| DR | | InterPro; IPR000209; Pept_58_553. |
| DR | | Pfam; PF02225; PA; 1. |
| DR | | Pfam; PF00082; Peptidease_S8; 1. |
| DR | | PRINTS; PRO0723; SUBTILISIN. |
| DR | | PROSITE; PS00137; SUBTYLAZE_HIS; UNKNOWN_1. |
| DR | | PROSITE; PS00138; SUBTYLAZE_SER; 1. |
| DR | | PROSITE; PS00138; SUBTYLAZE_SER; 1. |
| GO | | SEQUENCE 1294 AA; 135576 MW; 804C7F9ADDEE896 CEC64; |

| | | | | |
|-----------------------|------------------|--------------------|------------|--------------|
| Query Match | 21.9% | Score 752; | DB 2; | Length 1294; |
| Best Local Similarity | 45.1% | Pred. No. 1.3e-32; | | |
| Matches 173; | Conservative 62; | Mismatches 125; | Indels 24; | Gaps 9 |

[illegible]

| | | | |
|----|-----|--|-----|
| Db | 362 | LFVAAAGNSGPGSGTIGSPGADAAITGAVDRDLSLFFSSRGRSDDEAVRDPVTAPG | 421 |
| Qy | 351 | VDIAPRASGTSNGPFIINDYTTKASGTSMATPHYSGVGLILOAHPSTDPKVTALLET | 410 |
| Dd | 422 | VGIYAARATGTGMDPVDVDTGTTAAAGTSMATPHVAGAAALAAQRHPRMSAQLDVALVST | 481 |
| Qy | 411 | ADIVAPKEIADIALYGAQRVNVTKA | 434 |
| Dd | 482 | ARTIAGQVTE--QGGRIDILAA | 503 |

RESULT 6

Q9RL54 STROCO
ID Q9RL54 STROCO PRELIMINARY; PRT; 1245 AA.
AC Q9RL54;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Probable secreted peptidase.
GN OrderedLocustNames=SC00432; ORFNames=SCF51A.10;
OS Streptomyces coelicolor.
OC Streptomyces
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
NCBI_taxid=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AJ3(2) / M145;
RX MEDLINE:21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Klever H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornby T., Howarth S.,
RA Huang C.-H., Kießer T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabbittowitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor AJ3(2).";
TL Nature 417:141-147(2002).

CC
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
EMBL AL939105; CAB56662.1; - Genomic DNA.

DR HSSP; P00782; 1SUB.
DR BioCq; SCOE1902:SC0432-MONOMER; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR003137; PA.
DR InterPro; IPR000209; Pept_S8_553.
DR InterPro; IPR010221; VCBS.
DR Pfam; PF02225; PA; 1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISTIN.
DR TIGRfam; TIGR01965; VCBS_repeat; 1.
DR PROSITE; PS00137; SUBTILAS_HIS; UNKNOWN_1.
DR PROSITE; PS00138; SUBTILAS_SER; 1.

| | | | | | |
|-----------------------|--------------------|----------|------------------|------------------|--------------|
| AM | Complete proteome. | | | | |
| SEQ | SEQUENCE | 1245 AA; | 130896 MW; | 74EE92DB9CAlDE60 | CRC64; |
| Query Match | | 21.8%; | Score 748.5; | DB 2; | Length 1245; |
| Best Local Similarity | | 42.0%; | Pred. No. 2e-32; | | |
| Matches 181; | Conservative | 66; | Mismatches 125; | Indels 59; | Gaps 13 |

[illegible]


```

QY      177  OQKVIIGVNFVNRSPYPDHGHGHTFVASIAAGTAAANGKXKMAPEAKIAGIKVLGAD 236
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      252  AGRVAAAKDP--SSSGSTNDVFGHGHTHVASIVGSSGAHSGSSRGVAPPAKILVSKVLGDD 310
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      237  GSGSISTIIKGVEMAVDNKDKYGIKIVINLSLSSQSSTGDTLSLQAVNN--AMDAGIVCV 295
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      311  GFSSBSQVILAGHEMADQ----GADVNMSSLGSSSATTGTCTPMQALNDLSRRTGTLFVY 366
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      296  AAGNSG--PNTTYVGSPPAAASKVITGVAVDSNDNIASFSSRGP--TADGLKEVVAPEVDI 353
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      367  AAGNEGEQGPRTVGSPPGADALITGVADVDRDLSLAFSSRGPRLGDADVXDPVTAPEVGI 426
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      354  IAPRAGTSMKPIPIINDYTTKASGTSMATPHYSGVALLLOHPSTPDKXTALILEADI 413
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      427  VAPARAAGSAMGPVEBHYHTAASGTSMATPHVAGAAALLAOQHPPMTGNOILDALISIAVTT 486
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      414  VAPKEIADIAYAGAGRVNYKAIKYDDYAKLFTGSVADKGSATHTPDSGATFTATLWY 473
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      487  VDGQKYTE--QGGGRIDVRAA---GLGAVTATGTL-----VMGPFTSRDTEPVTSKRY 535
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      474  DTGSSDIDLVL 484
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      536  -TMSSEDEVTL 545
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 7

ID OSNM24_9ARCH PRELIMINARY; PRT; 795 AA.
AC OSNM24;
DT 04-JAN-2005, integrated into UniprotKB/TrEMBL.
DT 04-JAN-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Alkaline serine protease.
DE ORFNames=orf17;
GN uncultured archaeon.
OS uncultured archaeon.
OC Archaea; environmental samples.
NCBI TaxID=115547;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=16329940; DOI=10.1016/j.femsec.2004.12.004;
RA Ertel C, Kemnitz D, Kube M, Rieke P, Chin K.-J., Dedysh S.,
RT Reinhard R, Conrad R., Liesack W.,
RT Retrieval of first genome data for rice cluster I mechanisms by a
LT combination of cultivation and molecular techniques.";
LL FEMS Microbiol. Ecol. 53:187-204(2005).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC
DR EMBL, CR636956; CAHQ04782.1; - Genomic DNA.
DR GO:0008823; F:peptidase activity; IEA.
DR GO:0004280; F:protein self binding; IEA.
DR GO:0004289; F:subtilase activity; IEA.
DR GO:0004306; P:negative regulation of enzyme activity; IEA.
DR GO:0006508; P:proteolysis; IEA.
DR InterPro: IPR011964; Beta_ptc_yvnc.
DR InterPro: IPR000209; Pept_G6_S53.
DR InterPro: IPR010259; Prot_tmh_S8A.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF05922; Subtilisin_N; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR TIGRFAMs; TIGR02276; beta_ptc_yvnc; 3.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE; PS00138; SUBTILASE_SER; UNKNOWN_1.
DR KEGG
Q SEQUENCE 795 AA; 82723 MW; 629536E8B3D16BE8 CRC64;

| | | | | |
|---------------------------|-------|-------------------|------------|------------|
| Query Match | 21.5% | Score 737.5 | DB 2 | Length 795 |
| Best Local Similarity | 32.3% | Pred. No. 4.6e-32 | | |
| Matches 200, Conservative | 90 | Mismatches 208 | Indels 121 | Gaps 17 |

```

Qy 51 QKPNPNEEISTAYIVENHREKEIAVRVELMGAVRYVYHIIPIADLKRDLLVLSGL 110
Db 26 KKTIDNQGASLLSQGAASKEAKTRDLVKAGVKRYKNYNIIDAVADLPDSEVAAL--- 82
Qy 111 TGGKATLSGVRITQEDYKYVTSAELEGGDESAAOVMATYVNL--GYDSCITIGIIDTG 168
Db 83 ---KAR-PVSESIVRD-----SIAYADDEPVYHVNASVAMSTVPGYTGKGVNVSVIDSG 133
Qy 169 IDASHDLOGKYIGVWDVFNGBRSYVYDDHGHGTHVASIAGTGAASNGKYKGA PGAKLA 228
Db 134 IDHTPHDLGKYIIMKDPFLNDRATPYDDPGHGTVAGIIASGSGMSGEGYKGLAYDLSLF 193
Qy 229 GIKVTLGADSGSISTIIKGVEMAVDNKDXYGIKYIINLSLSSOSSDGTDSLISQAVNNAMD 288
Db 194 GYKVLNPSGTAVYSDIIAIDMSVQNH---ADVISMSLSNPTH---IQALDDAVHNAAE 246
Qy 289 AGIYVGVAAAGNSGPNITYVSGPAAASKITYTGAVDNSDNINASSRSRPTADGRKPEVVA 348
Db 247 NGVVVCSAGNTPKPKSGIRCPGSDPVIAGSVSMSSRPSFSSRBPDPGRKKPDI VA 306
Qy 349 PGVDIIAPASGSTMGPTRINDYYTASGSTMATPHVSGVALIIQAHSPWTDKVKALII 408
Db 307 VGEYVIVSYSSSGSTMGNEPQGYYTCYASGTSAAACQVSAASAILLQANGSLTPEIIRKOLI 366
Qy 409 ETADIAPAREIADI---AYGGRVNVYKAIKYDYAKLTFGTSVADKGSATHTFDVSGA 464
Db 367 RMT-----YHLSDTY PENOQGWGRINIA PAL--NEVLQVTPPTPTPTATPTPTPTPT 419
Qy 465 TPTATLALYDGTGSSDIDLXLYDPNGNEVDYSYTA YGEBKGYNNPNA--GWTYVAVVS 521
Db 420 ATPTAT-----PPTATPTPTPTPTPTPTATPTPTPTPTPTPTPTPTPTPTPTPTPT 440
Qy 522 YKGAANYQVADVDSGLSOSGGGNPNPNPNPTPT-----TDTOTFT. 564
Db 441 -----TYPTPTATPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 485
Qy 565 GSVNDYMDTSDTFTVN-----VNSGATKI TSD-LTFDTSYNDLXLVDPNGNLVDR 615
Db 486 PFTPTPTPVSKTFFVNSGNSNTVSVIDQSTNTVTGTIIITGSRPYG-----IVHSPDGSIV-- 539
Qy 616 STSSNSYEHVEYANPARGT 634
Db 540 YVAVEGLNRVALIISPANN 558

```

RESULT 8

| ID | OSENV1_OCEIH | PRELIMINARY; | PRT; | 430 AA. |
|--------------------|--|-----------------------------------|------|---------|
| AC | OSENV1_OCEIH | | | |
| DT | 01-MAR-2003 | integrated into UniProtKB/TrEMBL. | | |
| DT | 01-MAR-2003 | sequence version 1. | | |
| DT | 07-FEB-2006 | entry version 19. | | |
| DE | Intracellular alkaline serine proteinase. | | | |
| GN | OrderdedocusNames=OB3375; | | | |
| OS | Oceanobacillus theyensis. | | | |
| OC | Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus. | | | |
| NCBI_TaxID=182710; | | | | |

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=22220767; PubMed=12235376; DOI=10.1093/nar/gkE526;
RA Takami H., Takaki Y., Uchiyama I.;
RA "genome sequence of *Corynebacterium iheyensis* isolated from the Ihey
RT Ridge and its unexpected adaptive capabilities to extreme
RU environments.",
RL Nucleic Acids Res. 30:3927-3935(2002).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL, BA000026; BAC14331.1; -; Genomic_DNA.
DR HSSP; Q99405; IMPT.
DR BioCyc; Q1HE182710:OB2375-MONOMER; -;
CC

DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis; IEA.
 DR InterPro; IPR000209; Pept_S8_S53.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Complete proteome; Hydrolase; Protease; Serine protease.
 DR SEQUENCE 430 AA; 45839 MW; 6D09A95BCE1E310F CRC64;
 SQ
 Query Match 20.3%; Score 715.5; DB 2; Length 430;
 Best Local Similarity 46.8%; Pred. No. 3.4e-31;
 Matches 148; Conservative 54; Mismatches 95; Indels 19; Gaps 6;
 QY 138 LDESAQVMATYWNIGYDSSGITIGIDTGDASHPDLOGKIVGWDFNGSYPPDH 197
 DB 120 LPTASSINADVKEGSLTGQGSTIAVIDGHI-PHEDEGRITGFADFYKQTEPYDDN 178
 QY 198 GHGTHTVASTAGTGAANGKYKMGAPGAKLAGIKVLGADSSGSIPTTIKGVMAVDNKK 257
 DB 179 GHGTHTVAGDAGAGALSDGYQGPAPANLVGVKLNKTSGLSTVIBGIDMCIONSK 238
 QY 258 YGKIVNLISGS--SQSSDGTDLISQAVNNAMPAGIVCVAAAGSGPNTYVGSPPAASK 315
 DB 239 YNNINILSLSGSATPAEG-DEVNAVERFAMNGVVCVAAAGSGPDKTVGSPGISPK 297
 QY 316 VITVGAVDN-----DNIAFSRGRPTADGRLKBEVAVPVDTIAPRAG-----TSM 363
 DB 298 VITVGAADDNNTAERSDSSVAEFSRGPITDGLTKPNLLTPGVDIYSLRAPSFDIKTNK 357
 QY 364 GTFINYYTKASGSMATPHVSGVALILQAHPSMTPDXYKTALLETADIVAKETADIA 423
 DB 358 SARVGSNYISLSGTSMAETPCAGIVAOLOSLSLTPNVKEXKMEACODLGO---SPNV 414
 QY 424 YGAGRVVYVYKIKYDD 439
 DB 415 QGAGVYLAANLININE 430
 RESULT 9
 Q82139 STRAW PRELIMINARY; PRT; 1139 AA.
 ID Q82139 STRAW
 AC Q82139
 DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 16.
 DE Putative subtilisin-like protease.
 GN Ordered locus names=SAV3319;
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 microorganism Streptomyces avermitilis";
 RT Nat. Biotechnol. 21:526-531(2003).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 avermitilis: deducing the ability of producing secondary
 metabolites";

RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 CC -----
 CC Copyright by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonCommercial License
 CC -----
 DR EMBL; BA000030; BAC71030.1; -, Genomic_DNA.
 DR HSSP; Q99405; 1MPT.
 DR MEROPS; S08.069; -.
 DR BioCyc; SAV3319-MONOMER; -.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis; IEA.
 DR InterPro; IPR002860; Glyco_Hydro_BNR.
 DR InterPro; IPR000209; Pept_S8_S53.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Complete proteome; Protease.
 DR SEQUENCE 1139 AA; 118193 MW; EC3BD234FA94FA0 CRC64;
 SQ
 Query Match 20.5%; Score 703; DB 2; Length 1139;
 Best Local Similarity 34.3%; Pred. No. 5.6e-30;
 Matches 205; Conservative 80; Mismatches 215; Indels 98; Gaps 22;
 QY 110 LITGKRAKSGVRFIDQYKVTYSAELEGLDESAQVMATYWNIGYDSSGITIGIDTGI 169
 DB 210 VITNGDVTASGVAVHWLD-----GTRKASLDKSPQIGAPLMAAGDGGKIVAVLDTV 264
 QY 170 DASHPDLQKIVGWDFVNGRSYPPDDHGHGTHTVASIAAGTGAANGKYKMGAPGAKLAG 229
 DB 265 DATHPDLKQVAVESKNF-SAAADADHPGHGTHTVASIAAGTGAANGKYKMGAPGAKLAG 323
 QY 230 IKVLGADSSGSIPTTIKGVMAVDNKKYGIKYNISLSGSQSSDGTDLISQAVNN-AMD 288
 DB 324 GKVLDDTSGSGDSGLIAGEMAAEQ-----GADVNNISLGGCDLPE-IDPLEAEVNLSEB 378
 QY 289 AGIVCVAAAGNSP-NTYVGSPPAASKVITVGAVDNNDNISFSSRGPPTADRLKEPV 347
 DB 379 KGIPLAAGNEEFGEQITGSPGSAADALTVGAVDSDKLASFSSRGFLDGLAIKPDVT 438
 QY 348 APGVDTIAPRAGSGTSGTPIIN--DYTKASGSMATPHVSGVALILQAHPSMTPDXYK 404
 DB 439 APGVDTIAPRAGSVLDQGVQKPDGLYLTISGSMATPHVAGAAALILKQHPWMSFAELK 498
 QY 405 TALLETADIVAPREIADIAVYAGRVNVYAIKYDDYAKLFTFGSVADKGSAT-----H 457
 DB 499 GALTGSA--KGGKYTPFOGSGRIADVRAIK-----OSVINPNVSFPGIQOWPH 546
 QY 458 TPDVSGATVTYNTL-YMDGSSSDIDLXL-----YDPNGNEVDYSYTYVGPGEKYGYNPTA 512
 DB 547 TDDKP---VTQQLYRNLSGTSVTLNLASTATNPG-----VAAPSGFFKLGATKTV 556
 QY 513 GTWTVKVSYKGAANYQVVDVSDGSL-----OSGGNP-----N 547
 DB 597 PAGRKASVDF--VYNKLGTTDGAISAVYATGGQGYRTAAVQREVSYDVLKHD 654
 QY 548 PNPENPPTPTDTQTFTGSVNDV---DTSDTFTNNVNSGATKLTGDLTFD---TSYN 559
 DB 655 RDKGPVNVSTDLTGVSGLAADKMFAPYDASGTVKRVYKGNFLINASLFADEDEFTKGA 714
 QY 600 DLDLYYDNGNLVDSST---SSNSVHYEYANPAPAGTMTFLVYAYSTGMADYOLKA 654
 DB 715 D---WIAQPLSVTKNTTVVDARKAKPVDITVPDKG-----AKSAFASPDYVEA 762
 RESULT 10
 Q8KXK6 STRVD
 ID Q8KXK6 STRVD PRELIMINARY; PRT; 1105 AA.
 AC Q8KXK6
 DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.

DT 01-OCT-2002, sequence version 1.
 DE 07-FEB-2006, entry version 16.
 DT 1,4-dihydropyridine enantioselective esterase precursor.
 GN Name=dnpA;
 OS Streptomyces viridosporus.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomyces; Streptomyces; Streptomyces.
 NCBI_TaxId=67581;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=A-914;
 RX MEDLINE=2034940; PubMed=12039725;
 RX DOI=10.1128/AEM.68.6.2716-2725.2002;
 RA Aritaoka T., Matsufuji M., Nakashima T., Dobashi K.,
 RA Yoshikawa T., Yamada S., Momose H., Taguchi S.;
 RT "Streptomyces serine protease (DHP-A) as a new biocatalyst capable of
 RT forming chiral intermediates of 1,4-dihydropyridine calcium
 RT antagonists.";
 RL Appl. Microbiol. 68:2716-2725(2002).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC EMBL; AB007809; BAC00500.1; -; Genomic_DNA.
 DR HSSP; P00782; 2SER.
 DR MEROPS; S08.069; -.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis; IEA.
 DR InterPro; IPR000209; Pept_S8_S53.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR HydroLase; Protease; Serine protease; Signal.
 FT SIGNAL 1 204 Potential.
 SQ SEQUENCE 1105 AA; 114126 MW; 6AB34DB8B70579AB CRC64;
 Query Match 20.5%; Score 702; DB 2; Length 1105;
 Best Local Similarity 33.1%; Pred. No. 6.2e-30;
 Matches 205; Conservative 63; Mismatches 227; Indels 124; Gaps 22;
 QY 110 LRGKAKLGSVRIQDQDKYKTALEGLDESAQVMATYVNLGVDGSGITIGIIDTG 169
 DB 182 VYNGDRTAGIAHWLD-----GVRRALDTSVGOIGAPKAWAGYDGKGVKLAVIDTG 236
 QY 170 DASHPLQGVKVGWVDFVNGRSYFYDDHGHVNASIAAGTGAASNGKYKMAAGAKLAG 229
 DB 237 DTSHPDLKGRVTKAFTAPG-AGDKVGHGHVNASIAAGTGAASNGKYKMAAGAKLAG 295
 QY 230 IKVLGADGSGSISTIIKGVEMAVDNKRGYIKVINDLSGSSGSSDGTSLGQAVNN-AMD 288
 DB 296 GKVLDSGFGDDBGLAGHMAA---AQADAVNNSLGMDTPR-TDPLEAAVDKLSAE 350
 QY 289 AGIVVVAAGNSPNTYTGSPAAASKVTVGVAVDSNDNINASSSGRP-TADRLKPEVY 347
 DB 351 KGVLFIAAGNEPPE--SIGSPSADALVGVAVDKDLADPSISGPRGDAIKRPDVT 408
 QY 348 APGVDDIAPRASGSGTPIND---YYTKASGTSMPHYSGVGAIILOAHPGTPDKYK 404
 DB 409 APGVDDITTAASAEENDIGQEVGSEPPAGYMTISGTSMTPHVGAALILKQHPMTSAELK 468
 QY 405 TALIEADIAVPEIADIAVAGRVNVAIK----- 436
 DB 469 GAL--TGSTKGK-YPPEFGSGRIQADKALQOTVIADPVSVFGVQVQWPHTDDEPVTK 525
 QY 437 -----YDYAKLFTFTGSVADKSGATHTFDVSGATFYATLYMPDGSSDIDLYLDPN 488
 DB 526 LITRNIGTQDVTIKLSTATDPKGAAPAGFFTLGATTVVPA---GGSSAVMTADTRL 582
 QY 489 GNEVDYSYTAAYGFEKGVYVNPAGTWTVCVSY-----KGAANYQVDVVSDG 537

DB 583 GGTVDGASVAYVATGGGQVTRTAAAVQREVESYDVTRVHGRDGKPTTEHLTDLIGVAG 642
 QY 538 LSQSGGNGNPNPNPNPPTTDTOTFTGSVNDY-----W-----D 572
 DB 643 LSSGRGTC-----APATDTATLRPKGYLVDSWIAKDPGTLKGIIDLVPKLSV 693
 QY 573 TSDFTFMVNSGATKITYGLDFE-DTSYNDLDE---YLDPNGNLVDRSTSSNSYEHVEYA 628
 DB 694 TKDT-TLTLIDARTK-ADITVPDPKAPKLSATIGTYVTAG--IGIGASMSFADVRMA 749
 QY 629 NPAP-----GTWT 636
 DB 750 HIGPEAPGRIQRTWNGQWT 768
 RESULT 11
 ID Q651P4_BACLD PRELIMINARY; PRT; 442 AA.
 AC Q651P4; Q62U51;
 DT 25-OCT-2004, integrated into UniProtKB/TrEMBL.
 DT 25-OCT-2004, sequence version 1.
 DT 07-FEB-2006, entry version 15.
 DE AprX (Alkaline serine protease).
 GN Name=aprx; OrderedLocustNames=BL00903, BL102185;
 OS Bacillus licheniformis (strain DSM 13 / ATCC 14580).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI_TaxId=279010;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=1533718; DOI=10.1155/000079829;
 RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
 RA Ehrenreich P., Baumer S., Henne A., Liesegang H., Mekl R.,
 RA Ehrenreich A., Gottschalk G.;
 RT "The complete genome sequence of Bacillus licheniformis DSM13, an
 RT organism with great industrial potential."
 RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15461803; DOI=10.1186/gb-2004-5-10-r77;
 RA Key M.W., Kamalya P., Nelson B.A., Brody-Karpi S.D., Zaretsky E.J.,
 RA Tang M.W., Lopez de Leon A., Xiang H., Gueti V., Clausen I.G.,
 RA Olsen P.B., Raemussen M.D., Andersen J.T., Joergensen P.L.,
 RA Larsen T.S., Sorokin A., Bolotin A., Lapidus A., Galleron N.,
 RA Ehrlich S.D., Berka R.M.;
 RT "Complete genome sequence of the industrial bacterium Bacillus
 RT licheniformis and comparisons with closely related Bacillus species.";
 RL Genome Biol. 5:RESEARCH077.1-RESEARCH077.12(2004).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC EMBL; AB017333; AAU41070.1; -; Genomic_DNA.
 DR EMBL; CP000002; AAU23708.1; -; Genomic_DNA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis; IEA.
 DR InterPro; IPR000209; Pept_S8_S53.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN_1.
 DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Complete proteome; Protease.
 SQ SEQUENCE 442 AA; 48227 MW; 834B9CC27DFD83A CRC64;
 Query Match 20.1%; Score 688; DB 2; Length 442;
 Best Local Similarity 37.9%; Pred. No. 1.2e-29;
 Matches 158; Conservative 86; Mismatches 135; Indels 38; Gaps 11;
 QY 45 GLPKTKQANPBEISTVVPENREKEIAVAVLELMA---KRYVYHIIPALADLK 100
 DB 41 GFEKEI--LMKKKIPVILIEFKDCQHKGYOLVNEIGSKRNKIKHRSNVS CCSAE-- 96

QY 101 VRDLVITGSLTGKRAKLSGVRFIQEDYVTVSALEBGLDSEAAQVMATYVNLGYDGGSI 160
 Db 97 ---VTPSLQSLISECGDIRKRYLNKRYKALLVAVSSSHAKEVERN---NQTLTGKV 149
 QY 161 TIGIITDIDASHPDLQGVWDFVNGRSPYVDHGHGTHVSIAGTGAANGCYKG 220
 Db 150 TVAVITDGV-YPHEDLEBRIFAPODFINORKEPDDHGHCHGAGDAGNAAASGGYRG 208
 QY 221 MAPGAKIAGIVLGAADGSGSISTTIKGYEMAV---DNKQKYGKVINLSLSSGSS---QSS 273
 Db 209 PAPEAEIVGVKVLKMGSGSLFETVQGVDMCIQFNKENPDP-PDIIISMGLAALAYEN 267
 QY 274 DGTSLGAVANNAMDAGIVCVAAAGSGPNTYTGSPAAASKVITVGAVD-----SND 326
 Db 268 EEDPVPVAAVAAWDAGIVCVAAAGNSGPDQITIASPVSISKITVVALDPRDVTGRED 327
 QY 327 NIASFSSKGPADRLKPEVAVPGVDIIAPASGTSMG-----TPIDYTYKASGTSMAT 381
 Db 328 DVASYSRSGPTIYQGVWDFVNGRSPYVDHGHGTHVSIAGTGAANGCYKG 220
 QY 382 PHSVGVALIIQAHPSWTPDKVKTALITETADIVAPKEIADIYAGAGRVNYKAIKYD 438
 Db 388 PICAGIALIIOQAPGTEPDEVKQLMDGTDLM--KDRDPRVYGAGVYINAGSVPOD 442

RESULT 12

031788_BACSU PRELIMINARY; PRT; 442 AA.
 ID 031788_BACSU PRELIMINARY; PRT; 442 AA.
 AC 031788;
 DT 01-JAN-1998, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 1.
 DE Alkaline serine protease.
 GN Name=apix; Ordered locus names=BSU17260;
 OS *Bacillus subtilis*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
 RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
 Azevedo V., Besterio M.G., Bessieres P., Bolochin A., Borchert S.,
 Boursier L., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 Brouiller S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 Choi S.-K., Codani J.-J., Conerton I.F., Cummings N.J., Daniel R.A.,
 Denicot F., Devine K.M., Duesterhoeft A., Ehrlich S.D., Emerson P.T.,
 Ertlan K.-D., Errington J., Fabret C., Ferrati E., Foulger D.,
 Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 Gim S.-Y., Glaser P., Goffeau A., Golightly E.U., Grandi G.,
 Giuseppe G., Guy B.J., Hage K., Hache J., Harwood C.R., Henaut A.,
 Hilbert H., Holsappel S., Hosono S., Hullo M.-F., Itaya M.,
 Jones L.-M., Jorle B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
 Klein C., Kobayashi Y., Koetter P., Koningstein G., Krogh S.,
 Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
 Lazarevic V., Lee S.-M., Levine A., Liu H., Masuda S., Maue C.,
 Medigue C., Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S.,
 Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
 Park S.-H., Patro V., Pohl T.M., Portelle D., Portolillo S.,
 Prescott A.M., Pressac E., Pujic P., Purnelle B., Rapoport G.,
 Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
 Rose M., Sadate Y., Sato T., Scanlan R., Schleich S., Schroeder R.,
 Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serrit P.,
 Shin B.-S., Solido B., Sorokin A., Taccoti E., Takagi T., Takahashi H.,
 Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P.,
 Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vanler F.,
 Vasseroiti A., Viari A., Wambuit R., Wedler B., Wedler H.,
 Weitzengger T., Winters P., Wipac A., Yamamoto H., Yaneane K.,
 Yasumoto K., Yata K., Yoshida K., Yoshikawa H.-F., Zumbstein E.,
 Yoshikawa H., Zanchin A.,
 RA "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*.";
 RL Nature 390:249-256(1997).

CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC
 DR EMBL: Z99113; CAB13610.1; -; Genomic DNA.
 DR PIR: A69587; A69587.
 DR HSSP: Q99405; 1MP.
 DR BIOCYC: BSUB1423; BSU1727-MONOMER; -.
 DR GO: GO:0008233; F:peptidase activity; IEA.
 DR GO: GO:0004289; F:subtilase activity; IEA.
 DR GO: GO:0006508; F:proteolysis; IEA.
 DR InterPro: IPR00209; Pept_S8_S53.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KW Complete proteome; Hydrolase; Protease; Serine protease.
 SQ SEQUENCE 442 AA; 47906 MW; B996446AB87BADF2 CRC64;

Query Match 20.1%; Score 688; DB 2; Length 442;
 Best Local Similarity 39.7%; Pred. No. 1,2e-29;

Matches 167; Conservative 78; Mismatches 132; Indels 44; Gaps 14;

QY 43 TGLPFRKI--QKLNPEISITVIVENHREK--EIAVRVLEL-MGAKRVYVYIIIPAIA 96
 Db 35 TPCFLHKFPETKLNRRKMSVIEFEBCHEGTFQMGASVILQKRSKLSKRFNNINCS 94
 QY 97 ADLKRDLVITGSLTGKRAKLSGVRFIQEDYVTVSALEBGLDSEAAQVM-ATYVNLG- 154
 Db 95 AR-----VTPSLQSLISECGDIRKRYLNKRYKALLVAVSSSHAKEVERN---NQTLTGKV 142
 QY 155 -YDGSGITIGIITDIDASHPDLQGVWDFVNGRSPYVDHGHGTHVSIAGTGAANGCYKG 220
 Db 143 TLTGKVTVAVVITDGV-YPHEDLEBRIFAPODFINORKEPDDHGHCHGAGDAGNAAASGGYRG 208
 QY 214 SNGYKGAAPGAKIAGIVLGAADGSGSISTTIKGYEMAV---DNKQKYGKVINLSLSSGSS 269
 Db 202 SSGYRGPAPPEANLIGVKNKQSGTLADIIEGVEWCIQYNEPDE-PIDIMSMSSIGG 260
 QY 270 S---QSSGCTDLSGAVANNAMDAGIVCVAAAGNSGPDQITIASPVSISKITVVALDPRDVTGRED 327
 Db 261 DALRYDHEQDEPLVAVEAMSAAGIVCVAAAGNSGPDQITIASPVSISKITVVALDPRDVTGRED 327
 QY 326 -----DNIAFSSKGPADRLKPEVAVPGVDIIAPASGTSMG-----TPIDYTYKASGTSMAT 381
 Db 321 TASSDDVIVAFSSSGPTIYQGVWDFVNGRSPYVDHGHGTHVSIAGTGAANGCYKG 220
 QY 375 SGTSMATPHVSGVALIIQAHPSWTPDKVKTALITETADIVAPKEIADIYAGAGRVNYKAIKYD 438
 Db 381 SGTSMATPICAGIALIIOQAPGTEPDEVKQLMDGTDLM--KDRDPRVYGAGVYINAGSVPOD 442
 QY 435 I 435
 Db 439 V 439

RESULT 13

051315_GEOKA PRELIMINARY; PRT; 442 AA.
 ID 051315_GEOKA PRELIMINARY; PRT; 442 AA.
 AC 051315;
 DT 01-FEB-2005, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 1.
 DE Intracellular alkaline serine protease.
 GN Ordered locus names=GK0210;
 OS *Geobacillus kaustophilus*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
 OX NCBI_TaxID=1462;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=HTA426;
 RX PubMed=1576355; DOI=10.1093/nar/gkh970;

CC Streptomycetaceae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1887;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=S-1253
 RX MEDLINE=97144528; PubMed=8990295;
 RA Suzuki M., Taguchi S., Yamada S., Kojima S., Momose H.;
 RT "A novel member of the subtilisin-like protease family from
 Streptomyces albobacillus.";
 RL J. Bacteriol. 179:430-438(1997).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 CC EMBL: D83672; BA12040.1; -; Genomic_DNA.
 DR HSSP; P00782; 2SBT.
 DR MEROPS; S08_069; -.
 DR GO: GO:0008233; F:peptidase activity; IEA.
 DR GO: GO:0004289; F:subtilase activity; IEA.
 DR GO: GO:0006508; P:proteolysis; IEA.
 DR InterPro; IPR002860; Glyco_hydro_BNR.
 DR InterPro; IPR00209; Pept_S8_S53.
 DR Pfam; PF00212; BNR; 2.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISTN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Protease.
 KM
 SQ SEQUENCE 1102 AA; 114128 MW; F9E4AD2590FE559E CRC64;

Query Match 19.5%; Score 668.5; DB 2; Length 1102;
 Best Local Similarity 32.2%; Pred. No. 4.3e-28;
 Matches 202; Conservative 71; Mismatches 228; Indels 127; Gaps 22;

QY 109 GLTGKAKSGVRFQEDYVTVSALEGLDESAQVMATYWNLTGDSGITIGITDG 168
 DB 172 GVTDGDRFASGVARWLD---GVRKASLDTSVGOIGPKAWAGYDGKGVKIAVLDTG 226
 QY 169 IDASHPDLQKVLGVWVFNVGRSPYDDHGHVHVASIAAGTGAASNGKYGMAFGAKLA 228
 DB 227 VDATHPDLKGVYASKNFTSAPT-TGVVGHGTHVASIAAGTGAOSKGYKGVAPGAKII 285
 QY 229 GIVYLADSGSGSTIIKGYEMAVDNKKYGIKVINLSGSSSQSDGTSLSQAVNN-AM 287
 DB 286 NGKVLDDAGFGDGGIAGMEMAA---AQGADIWNMSLGGMDTPE-TDPLEAAVDKLSA 340
 QY 288 DAGIVCVAAAGSGPNMTYTGSPAAASKVITVGAVDNDNIAFSRGP-TADGRLKEV 346
 DB 341 EKGILFRAIAGNBPQ--STGSPGASALTVGAVDDKDLADFSSTGPRLDGAVKPD 398
 QY 347 VAPGVDLIAFRASGTSGMTPIINDY---YTKASGTSMATPHVSGVALILOAHPMTPDY 403
 DB 399 TAPGVDTITASAKGNDAKEVGEKPGAGYMTISGTMATPHVGAALLKQHPMKYAEI 458
 QY 404 KTLA-IETADIVAPKIAIDIVAGRVNYKAIY-----DD--Y 440
 DB 459 KGLTLTSTKD---GKTPFEGSGRVQVDKAITQVIAPVSLSPGVQWPHADKPV 514
 QY 441 AKLTF---TGSVADKGSATHT-----FDVSGATFVATLYMDGSSDIDLXYLD 486
 DB 515 KULTYRNLTGEDVTLLKLTSTATGPKKAAPAGFTLGASTLVPA---NGTASVDVADT 571
 QY 487 PNGNEVDYSYATAYGFEKVGYNPTAGTWTVKVSY-----KGAANYQ----- 529
 DB 572 RLGAADVGTSAVAVTGAQGSVRTAAVERREVESYNTVLKVLDRSGKATANYMAYLSGL 631
 QY 530 -----VDVVSQSLG---OSGG-----GNPNPNPNPT 555
 DB 632 TGIKDRSYAPYEDAVSVRPKGGYVLDAVLYGADPETWGADWMLAQPKLDVTRNTT 691
 QY 556 PTTDTQ-----TFTGSVNDYWDTSDTFTMANVSGATKITGDLTFPTDSYNDLDLYNDP 608

DB 692 VTVDAKRAKPVKTVPGKAKAKQFASADYTIETNDSAVYGMWLENYSGFRSAHL----- 746
 QY 609 NGNLVDRSTSSNSNYEHVEYANPAPGTWT 636
 DB 747 -GPQITNGTLSQQW-NTHFSNGAKAQYT 772

Search completed: January 6, 2007, 22:31:22
 Job time : 197 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2007 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2007, 22:32:37 ; Search time 123.5 Seconds
(without alignments)
2471.729 Million cell updates/sec

Title: US-10-800-684-5
Perfect score: 3428
Sequence: 1 MKGKALIVIVLGLVVGVS.....YASTYGNADYOLKAVVYVG 659

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /EMC Celerra_SIDS3/Ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /EMC Celerra_SIDS3/Ptodata/2/pubpaa/US08_PUBCOMB.pep:*
3: /EMC Celerra_SIDS3/Ptodata/2/pubpaa/US09_PUBCOMB.pep:*
4: /EMC Celerra_SIDS3/Ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
5: /EMC Celerra_SIDS3/Ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
6: /EMC Celerra_SIDS3/Ptodata/2/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--|
| 1 | 3428 | 100.0 | 659 | 3 | US-09-841-553-5 Sequence 5, Appl1 |
| 2 | 3428 | 100.0 | 659 | 3 | US-10-800-684-5 Sequence 5, Appl1 |
| 3 | 2997.5 | 87.4 | 654 | 3 | US-09-841-553-35 Sequence 35, Appl1 |
| 4 | 2997.5 | 87.4 | 654 | 4 | US-10-090-624-16 Sequence 16, Appl1 |
| 5 | 2997.5 | 87.4 | 654 | 4 | US-10-800-684-35 Sequence 35, Appl1 |
| 6 | 2997.5 | 87.4 | 654 | 5 | US-10-888-588-16 Sequence 16, Appl1 |
| 7 | 2914 | 85.0 | 659 | 3 | US-09-841-553-1 Sequence 1, Appl1 |
| 8 | 2914 | 85.0 | 659 | 4 | US-10-090-624-12 Sequence 12, Appl1 |
| 9 | 2914 | 85.0 | 659 | 5 | US-10-800-684-1 Sequence 1, Appl1 |
| 10 | 2914 | 85.0 | 659 | 5 | US-10-888-588-12 Sequence 12, Appl1 |
| 11 | 2349.5 | 68.5 | 522 | 3 | US-09-841-553-3 Sequence 3, Appl1 |
| 12 | 2349.5 | 68.5 | 522 | 4 | US-10-090-624-4 Sequence 4, Appl1 |
| 13 | 2349.5 | 68.5 | 522 | 5 | US-10-800-684-3 Sequence 3, Appl1 |
| 14 | 2349.5 | 68.5 | 522 | 5 | US-10-888-588-4 Sequence 4, Appl1 |
| 15 | 1918 | 56.0 | 412 | 4 | US-10-090-624-1 Sequence 1, Appl1 |
| 16 | 1918 | 56.0 | 412 | 5 | US-10-888-588-1 Sequence 1, Appl1 |
| 17 | 703 | 20.5 | 1139 | 4 | US-10-156-761-10856 Sequence 10856, A |
| 18 | 662.5 | 19.3 | 1079 | 4 | US-10-112-488-39 Sequence 39, Appl1 |
| 19 | 662.5 | 19.3 | 1079 | 4 | US-10-673-860-8 Sequence 8, Appl1 |
| 20 | 662.5 | 19.3 | 1208 | 4 | US-10-156-761-13251 Sequence 13251, A |
| 21 | 655 | 19.1 | 1237 | 5 | US-10-314-657-4 Sequence 4, Appl1 |
| 22 | 655 | 19.1 | 1237 | 5 | US-10-473-193-4 Sequence 4, Appl1 |
| 23 | 562 | 16.4 | 379 | 3 | US-09-813-408-11 Sequence 11, Appl1 |
| 24 | 561 | 16.4 | 372 | 3 | US-09-813-408-9 Sequence 9, Appl1 |
| 25 | 560 | 16.3 | 379 | 6 | US-11-156-062-23 Sequence 23, Appl1 |
| 26 | 557 | 16.2 | 379 | 3 | US-09-920-118-14 Sequence 14, Appl1 |
| 27 | 557 | 16.2 | 379 | 4 | US-10-202-339-2 Sequence 2, Appl1 |

| | | | | | |
|----|-------|------|-----|---|---|
| 28 | 555 | 16.2 | 382 | 4 | US-10-090-624-31 Sequence 31, Appl1 |
| 29 | 555 | 16.2 | 382 | 4 | US-10-104-693-2 Sequence 2, Appl1 |
| 30 | 555 | 16.2 | 382 | 5 | US-10-800-684-45 Sequence 45, Appl1 |
| 31 | 555 | 16.2 | 382 | 5 | US-10-888-588-31 Sequence 31, Appl1 |
| 32 | 555 | 16.2 | 382 | 5 | US-10-498-714A-2 Sequence 2, Appl1 |
| 33 | 555 | 16.2 | 382 | 6 | US-11-127-727-2 Sequence 2, Appl1 |
| 34 | 555 | 16.2 | 382 | 6 | US-11-127-947-2 Sequence 2, Appl1 |
| 35 | 554 | 16.2 | 379 | 3 | US-09-813-408-10 Sequence 10, Appl1 |
| 36 | 554 | 16.2 | 379 | 4 | US-10-146-905A-10 Sequence 10, Appl1 |
| 37 | 554 | 16.2 | 379 | 5 | US-10-503-706-1 Sequence 1, Appl1 |
| 38 | 553 | 16.1 | 380 | 4 | US-10-146-905A-12 Sequence 12, Appl1 |
| 39 | 551 | 16.1 | 379 | 3 | US-09-813-408-13 Sequence 13, Appl1 |
| 40 | 551 | 16.1 | 382 | 4 | US-10-146-905A-8 Sequence 8, Appl1 |
| 41 | 550 | 16.0 | 380 | 3 | US-09-813-408-19 Sequence 19, Appl1 |
| 42 | 549.5 | 16.0 | 378 | 3 | US-09-813-408-14 Sequence 14, Appl1 |
| 43 | 548.5 | 16.0 | 380 | 5 | US-10-836-959-2 Sequence 2, Appl1 |
| 44 | 548.5 | 16.0 | 380 | 5 | US-10-476-463-2 Sequence 2, Appl1 |
| 45 | 547 | 16.0 | 382 | 4 | US-10-033-325-2 Sequence 2, Appl1 |

ALIGNMENTS

RESULT 1
US-09-841-553-5
Sequence 5, Application US/09841553
Publication No. US20020086402A1
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
MORISHITA, Mio
YAMAMOTO, Katsuhiko
MITTA, Masanori
ASADA, Kiyozo
TSUNAWA, Susumu
KATO, Ikumoshin
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESS: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/841,553
FILING DATE: 24-Apr-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/894,818
FILING DATE: <Unknown>
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-Dec-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 5
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-841-553-5

Div. of Immunology

Query Match 100.0%; Score 3428; DB 3; Length 659;
Best Local Similarity 100.0%; Pred. No. 4,6e-234;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKGLKALILVILVLGLVGSVAAAPKKVEQVRNVEKNYGLLTPGLFRKIQKLNPNBEIS 60
DB 1 MKGLKALILVILVLGLVGSVAAAPKKVEQVRNVEKNYGLLTPGLFRKIQKLNPNBEIS 60
QY 61 TVIVFENHREKEIAVRVLELMGAKRVYVHHIIPALADLKVDLIVISGLTGKAKLSGV 120
DB 61 TVIVFENHREKEIAVRVLELMGAKRVYVHHIIPALADLKVDLIVISGLTGKAKLSGV 120
QY 121 RFIQEDYKTVSAELGLDESAQAQVMAATYVNLGYDGSGITIGIIDTGIDASHPDLOQKV 180
DB 121 RFIQEDYKTVSAELGLDESAQAQVMAATYVNLGYDGSGITIGIIDTGIDASHPDLOQKV 180
QY 181 IGVAVDFVNRGSYPYDDHGHGTHVASTAAGTGAASNGKYGMPGAKLAGIKVLGADGSGS 240
DB 181 IGVAVDFVNRGSYPYDDHGHGTHVASTAAGTGAASNGKYGMPGAKLAGIKVLGADGSGS 240
QY 241 ISTIIGVEMAVNDKDKYGIKVINLSLGSQSSDGTDSLQAANNMADAGIIVCVAAAGNS 300
DB 241 ISTIIGVEMAVNDKDKYGIKVINLSLGSQSSDGTDSLQAANNMADAGIIVCVAAAGNS 300
QY 301 GPNVTYVGSPPAASKVITVGAVDSDNDIASFSSRGPTADGRLEPEVVAQGVIIAPRASG 360
DB 301 GPNVTYVGSPPAASKVITVGAVDSDNDIASFSSRGPTADGRLEPEVVAQGVIIAPRASG 360
QY 361 TSMGTPINDYTTKASGSMATPHVSGVGLIIQAHPSMTPDKYKTLITFADIIVAPKEIA 420
DB 361 TSMGTPINDYTTKASGSMATPHVSGVGLIIQAHPSMTPDKYKTLITFADIIVAPKEIA 420
QY 421 DIAYGAGRNVVYKAIKYDDYAKLFTGSAVDKGSATHTFVSGATFVATLTYMDTSSDI 480
DB 421 DIAYGAGRNVVYKAIKYDDYAKLFTGSAVDKGSATHTFVSGATFVATLTYMDTSSDI 480
QY 481 DLYLDPNGNEVDYSTAYYGFPEKVGYNPTAGTWTVKVSYKGAANYQVDVSDSLSQ 540
DB 481 DLYLDPNGNEVDYSTAYYGFPEKVGYNPTAGTWTVKVSYKGAANYQVDVSDSLSQ 540
QY 541 SGGGNPNPNPNPPTTDTOTFTGSVNDYMDSDFTNNVNSGATKITGDLTFPDSYND 600
DB 541 SGGGNPNPNPNPPTTDTOTFTGSVNDYMDSDFTNNVNSGATKITGDLTFPDSYND 600
QY 601 LDLYLDPNGNLVDRSTSSNSYEHEVYANPAPGTWTFVLYASTYGMADYQLKAVVYTG 659
DB 601 LDLYLDPNGNLVDRSTSSNSYEHEVYANPAPGTWTFVLYASTYGMADYQLKAVVYTG 659

RESULT 2
US-10-800-684-5
Sequence 5, Application US/10800684
Publication No. US20050014221A1

GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MITTA, Masanori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu

APPLICANT: KATO, Ikunobu
TITLE OF INVENTION: ULTATHERMOSTABLE PROTEASE GENES
FILE REFERENCES: TAKAKURA-1A
CURRENT APPLICATION NUMBER: US/10/800,684
CURRENT FILING DATE: 2004-03-16
PRIOR APPLICATION NUMBER: US/09/841,553
PRIOR FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: 08/894,818
PRIOR FILING DATE: 1997-08-29
PRIOR APPLICATION NUMBER: JP32385/1995
PRIOR FILING DATE: 1995-12-12
PRIOR APPLICATION NUMBER: JP96/03253

PRIOR FILING DATE: 1996-11-07
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn version 3.2
SEQ ID NO 5
LENGTH: 659
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-10-800-684-5

Query Match 100.0%; Score 3428; DB 5; Length 659;
Best Local Similarity 100.0%; Pred. No. 4,6e-234;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKGLKALILVILVLGLVGSVAAAPKKVEQVRNVEKNYGLLTPGLFRKIQKLNPNBEIS 60
DB 1 MKGLKALILVILVLGLVGSVAAAPKKVEQVRNVEKNYGLLTPGLFRKIQKLNPNBEIS 60
QY 61 TVIVFENHREKEIAVRVLELMGAKRVYVHHIIPALADLKVDLIVISGLTGKAKLSGV 120
DB 61 TVIVFENHREKEIAVRVLELMGAKRVYVHHIIPALADLKVDLIVISGLTGKAKLSGV 120
QY 121 RFIQEDYKTVSAELGLDESAQAQVMAATYVNLGYDGSGITIGIIDTGIDASHPDLOQKV 180
DB 121 RFIQEDYKTVSAELGLDESAQAQVMAATYVNLGYDGSGITIGIIDTGIDASHPDLOQKV 180
QY 181 IGVAVDFVNRGSYPYDDHGHGTHVASTAAGTGAASNGKYGMPGAKLAGIKVLGADGSGS 240
DB 181 IGVAVDFVNRGSYPYDDHGHGTHVASTAAGTGAASNGKYGMPGAKLAGIKVLGADGSGS 240
QY 241 ISTIIGVEMAVNDKDKYGIKVINLSLGSQSSDGTDSLQAANNMADAGIIVCVAAAGNS 300
DB 241 ISTIIGVEMAVNDKDKYGIKVINLSLGSQSSDGTDSLQAANNMADAGIIVCVAAAGNS 300
QY 301 GPNVTYVGSPPAASKVITVGAVDSDNDIASFSSRGPTADGRLEPEVVAQGVIIAPRASG 360
DB 301 GPNVTYVGSPPAASKVITVGAVDSDNDIASFSSRGPTADGRLEPEVVAQGVIIAPRASG 360
QY 361 TSMGTPINDYTTKASGSMATPHVSGVGLIIQAHPSMTPDKYKTLITFADIIVAPKEIA 420
DB 361 TSMGTPINDYTTKASGSMATPHVSGVGLIIQAHPSMTPDKYKTLITFADIIVAPKEIA 420
QY 421 DIAYGAGRNVVYKAIKYDDYAKLFTGSAVDKGSATHTFVSGATFVATLTYMDTSSDI 480
DB 421 DIAYGAGRNVVYKAIKYDDYAKLFTGSAVDKGSATHTFVSGATFVATLTYMDTSSDI 480
QY 481 DLYLDPNGNEVDYSTAYYGFPEKVGYNPTAGTWTVKVSYKGAANYQVDVSDSLSQ 540
DB 481 DLYLDPNGNEVDYSTAYYGFPEKVGYNPTAGTWTVKVSYKGAANYQVDVSDSLSQ 540
QY 541 SGGGNPNPNPNPPTTDTOTFTGSVNDYMDSDFTNNVNSGATKITGDLTFPDSYND 600
DB 541 SGGGNPNPNPNPPTTDTOTFTGSVNDYMDSDFTNNVNSGATKITGDLTFPDSYND 600
QY 601 LDLYLDPNGNLVDRSTSSNSYEHEVYANPAPGTWTFVLYASTYGMADYQLKAVVYTG 659
DB 601 LDLYLDPNGNLVDRSTSSNSYEHEVYANPAPGTWTFVLYASTYGMADYQLKAVVYTG 659

RESULT 3
US-09-841-553-35
Sequence 35, Application US/09841553
Publication No. US20020086402A1
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MITTA, Masanori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
APPLICANT: KATO, Ikunobu

TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES

NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/841,553
FILING DATE: 24-Apr-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/894,818
FILING DATE: <Unknown>
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-841-553-35

Query Match 87.4%; Score 2997.5; DB 3; Length 654;
Best Local Similarity 88.3%; Pred. No. 1.5e-203;
Matches 582; Conservative 22; Mismatches 50; Indels 5; Gaps 1;

1 MGGLKALLIIVILVGLGVSVAAAPKKYEQVNVKNGKGLTPGLFRKIQKLNPEEIS 60
1 MGGLKALLIIVILVGLGVSVAAAPKKYEQVNVKNGKGLTPGLFRKIQKLNPEEIS 60
61 TVIVFENHREKEIAVRLVLEMGAKRYVYHIIIPALADLKVRDLVITSGLTGSKATLSGV 120
61 TVIVFENHREKEIAVRLVLEMGAKRYVYHIIIPALADLKVRDLVITSGLTGSKATLSGV 120
121 RFIQEDYKVTVALEGLDESAQAQVMAVYVMNLGYDGSGITIGIIDTGIDASHPDLOGKY 180
121 RFIQEDYKVTVALEGLDESAQAQVMAVYVMNLGYDGSGITIGIIDTGIDASHPDLOGKY 180
181 IGVWDFVNRASYPYDDHGHGTHVASIAAGTGAASNGKYGMAFGALAIKYLGAAGSGS 240
181 IGVWDFVNRASYPYDDHGHGTHVASIAAGTGAASNGKYGMAFGALAIKYLGAAGSGS 240
181 IGVWDFVNRASYPYDDHGHGTHVASIAAGTGAASNGKYGMAFGALAIKYLGAAGSGS 240
241 ISTIIKGVMAVDNKKYGIKVINLSLSSQSSDGTDSLQAVNNAMDAIIVVCVAAGNS 300
241 ISTIIKGVMAVDNKKYGIKVINLSLSSQSSDGTDSLQAVNNAMDAIIVVCVAAGNS 300
241 ISTIIKGVMAVDNKKYGIKVINLSLSSQSSDGTDSLQAVNNAMDAIIVVCVAAGNS 300
301 GPNRTYVGSPPAASAKVITVGAVDSDNINIASFSRSGPTADGRLEPVVAPGMMIIAPRAG 360
301 GPNRTYVGSPPAASAKVITVGAVDSDNINIASFSRSGPTADGRLEPVVAPGMMIIAPRAG 360
361 TSMGPIINDYYTKASGTSMAFPHVSGVGLIIQAHPSWTPDKVKTALIIETADIAPKEIA 420
361 TSMGPIINDYYTKASGTSMAFPHVSGVGLIIQAHPSWTPDKVKTALIIETADIAPKEIA 420
421 DIAYGAGRNVVYKAIKYDDYAKLTFTGSAVDKGSATHTPDVSGATFVTATLVYDWTGSSDI 480
421 DIAYGAGRNVVYKAIKYDDYAKLTFTGSAVDKGSATHTPDVSGATFVTATLVYDWTGSSDI 480

481 DLVLYDPNGNEVDYSTAYATYGEKGYNPNATGATWTKVYSKGAANYQVDVSDSLQ 540
481 DLVLYDPNGNEVDYSTAYATYGEKGYNPNATGATWTKVYSKGAANYQVDVSDSLQ 540
541 SGGGNPNPNPNPNPTTDTOTFTGSDVNDYMDTSDTFMNVNNSGATKIGDLTFDTSYND 600
541 SGGGNPNPNPNPNPTTDTOTFTGSDVNDYMDTSDTFMNVNNSGATKIGDLTFDTSYND 600
541 SGGGNPNPNPNPNPTTDTOTFTGSDVNDYMDTSDTFMNVNNSGATKIGDLTFDTSYND 600
601 LBLVLYDPNGNLDVDSSTNSYEHVYANAPAGTMTFLVYASTVGMDADYOLKAVVYVY 659
601 LBLVLYDPNGNLDVDSSTNSYEHVYANAPAGTMTFLVYASTVGMDADYOLKAVVYVY 659
659 LBLVLYDPNGNLDVDSSTNSYEHVYANAPAGTMTFLVYASTVGMDADYOLKAVVYVY 659

RESULT 4
US-10-090-624-16
Sequence 16, Application US/10090624
Publication No. US20020132335A1
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hi-karu
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOJO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Kunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA=6
CURRENT APPLICATION NUMBER: US/10/090,624
CURRENT FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: 09/445,472
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 16
LENGTH: 654
TYPE: PRT
ORGANISM: Pyrococcus furiosus
US-10-090-624-16

Query Match 87.4%; Score 2997.5; DB 4; Length 654;
Best Local Similarity 88.3%; Pred. No. 1.5e-203;
Matches 582; Conservative 22; Mismatches 50; Indels 5; Gaps 1;

1 MGGLKALLIIVILVGLGVSVAAAPKKYEQVNVKNGKGLTPGLFRKIQKLNPEEIS 60
1 MGGLKALLIIVILVGLGVSVAAAPKKYEQVNVKNGKGLTPGLFRKIQKLNPEEIS 60
61 TVIVFENHREKEIAVRLVLEMGAKRYVYHIIIPALADLKVRDLVITSGLTGSKATLSGV 120
61 TVIVFENHREKEIAVRLVLEMGAKRYVYHIIIPALADLKVRDLVITSGLTGSKATLSGV 120
121 RFIQEDYKVTVALEGLDESAQAQVMAVYVMNLGYDGSGITIGIIDTGIDASHPDLOGKY 180
121 RFIQEDYKVTVALEGLDESAQAQVMAVYVMNLGYDGSGITIGIIDTGIDASHPDLOGKY 180
181 IGVWDFVNRASYPYDDHGHGTHVASIAAGTGAASNGKYGMAFGALAIKYLGAAGSGS 240
181 IGVWDFVNRASYPYDDHGHGTHVASIAAGTGAASNGKYGMAFGALAIKYLGAAGSGS 240
181 IGVWDFVNRASYPYDDHGHGTHVASIAAGTGAASNGKYGMAFGALAIKYLGAAGSGS 240
241 ISTIIKGVMAVDNKKYGIKVINLSLSSQSSDGTDSLQAVNNAMDAIIVVCVAAGNS 300
241 ISTIIKGVMAVDNKKYGIKVINLSLSSQSSDGTDSLQAVNNAMDAIIVVCVAAGNS 300
241 ISTIIKGVMAVDNKKYGIKVINLSLSSQSSDGTDSLQAVNNAMDAIIVVCVAAGNS 300
301 GPNRTYVGSPPAASAKVITVGAVDSDNINIASFSRSGPTADGRLEPVVAPGMMIIAPRAG 360
301 GPNRTYVGSPPAASAKVITVGAVDSDNINIASFSRSGPTADGRLEPVVAPGMMIIAPRAG 360
361 TSMGPIINDYYTKASGTSMAFPHVSGVGLIIQAHPSWTPDKVKTALIIETADIAPKEIA 420
361 TSMGPIINDYYTKASGTSMAFPHVSGVGLIIQAHPSWTPDKVKTALIIETADIAPKEIA 420
421 DIAYGAGRNVVYKAIKYDDYAKLTFTGSAVDKGSATHTPDVSGATFVTATLVYDWTGSSDI 480
421 DIAYGAGRNVVYKAIKYDDYAKLTFTGSAVDKGSATHTPDVSGATFVTATLVYDWTGSSDI 480

[illegible]

RESULT 5

```

US-10-800-684-35
/ Sequence 35. Application US/10800664
/ Publication NO. US20050014221A1
/ GENERAL INFORMATION:
/ APPLICANT: TAKAKURA, Hikaru
/ APPLICANT: MORISHITA, Mio
/ APPLICANT: YAMAMOTO, Katsuhiko
/ APPLICANT: MITTA, Masanori
/ APPLICANT: ASADA, Kiyozo
/ APPLICANT: TSUNASAWA, Susumu
/ APPLICANT: KATO, Ikunobu
/ TITLE OF INVENTION: ULTRATHERMOSTABLE PROTEASE GENES
/ FILE REFERENCE: TAKAKURA-1A
/ CURRENT APPLICATION NUMBER: US/10/800,684
/ CURRENT FILING DATE: 2004-03-16
/ PRIOR APPLICATION NUMBER: US/09/841,553
/ PRIOR FILING DATE: 2001-04-24
/ PRIOR APPLICATION NUMBER: 08/894,818
/ PRIOR FILING DATE: 1997-08-29
/ PRIOR APPLICATION NUMBER: JP3385/1995
/ PRIOR FILING DATE: 1995-12-12
/ PRIOR APPLICATION NUMBER: JP96/03253
/ PRIOR FILING DATE: 1996-11-07
/ NUMBER OF SEQ ID NOS: 45
/ SOFTWARE: Patentin version 3.2
/ SEQ ID NO 35
/ LENGTH: 654
/ TYPE: PRF
/ ORGANISM: Pyrococcus furiosus
/ US-10-800-684-35

```

| | | | | |
|-------------|-------|---------------|-------|-------------|
| Query Match | 87.4% | Score 2997.5; | DB 5; | Length 654; |
|-------------|-------|---------------|-------|-------------|

best local similarity 88.3%; Pred. No. 1.5e-203;
Matches 582; Conservative 22; Mismatches 50; Indels 5; Gaps 1;

| | | | |
|----|-----|--|-----|
| QY | 1 | MGKALKLIVILVILGLVVGVSVAAPAEKKVEQVAVNKNVYGLTPGLPRKIQKLNPNHEEIS | 60 |
| Db | 1 | MGLAKLILIVILVLGLVVGVSVAAPAEKKVEQVAVNKNVYGLTPGLPRKIQKLNPNHEEIS | 60 |
| QY | 61 | TVIIVENHREKILAVRVLELNGAKRRVYTHIIPALADLKRDLVTSGLTGSKANLSEV | 120 |
| Db | 61 | TVIIVENHREKILAVRVLELNGAKRRVYTHIIPALADLKRDLVTSGLTGSKANLSEV | 120 |
| QY | 121 | RFIOEDYKVTWSAELEGLDESAACQMATVYWNLTVDGSGITIGIIPDGDASHPDLOQKY | 180 |
| Db | 121 | RFIOEDYKVTWSAELEGLDESAACQMATVYWNLTVDGSGITIGIIPDGDASHPDLOQKY | 180 |
| QY | 181 | IGWVDPVNGRSYFYDDHGHGHTVNASIAAGTGAASNGKYKGMAPAKIAGIKVIGAGGSGS | 240 |
| Db | 181 | IGWVDPVNGRSYFYDDHGHGHTVNASIAAGTGAASNGKYKGMAPAKIAGIKVIGAGGSGS | 240 |
| QY | 241 | ISTIIIGVEMAVDNKDKYGIKVINLSTLGSQSSDGTDSLQAVNANWADAGIIVCVAAAGNS | 300 |
| Db | 241 | ISTIIIGVEMAVDNKDKYGIKVINLSTLGSQSSDGTDALQAVNAADAGLVVVAAGNS | 300 |
| QY | 301 | GPNTTVGVSPPAAASKVITVGAVDNSNDNIASFSSSGPFLADRKLKPEVVAPEVDIIPARAG | 360 |

[illegible]

RESULT

US-10-888-588-16
Sequence 16 Application US/10888588
Publication No. US20050084934A1
GENERAL INFORMATION:
APPLICANT: TAKAKURA, HiKaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOTO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikinoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/10/888,588
CURRENT FILING DATE: 2004-07-12
PRIOR APPLICATION NUMBER: US/10/090,624
PRIOR FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: 03/445,472
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 16
LENGTH: 654
TYPE: PRT
ORGANISM: *Pyrococcus furiosus*
US-10-888-588-16

| | | | | |
|-----------------------|--------|--|---------------|------------|
| Query Match | 87.44% | Score 2997.5 | DB 5 | Length 654 |
| Best Local Similarity | 86.33% | Pred. No. 1.5e-203 | | |
| Matches | 582 | Conservative 22 | Mismatches 50 | Indels 5 |
| | | | Gaps | 1 |
| QY | 1 | MKGKALILVILVGLVVGSAAPAEKKEVEQVRYVKKYGLLTGELFKRIQKLNPEIS | 60 | |
| Db | 1 | MKGKALILVILVGLVVGSAAPAEKKEVEQVRYVKKYGLLTGELFKRIQKLNPEIS | 60 | |
| QY | 61 | TVIVFENHREKEIAVRVLELMGAVRVVYHIIIPAIADLKVRDLLVLSGLTGGAKLSGV | 120 | |
| Db | 61 | TVIVFENHREKEIAVRVLELMGAVRVVYHIIIPAIADLKVRDLLVLSGLTGGAKLSGV | 120 | |
| QY | 121 | RFIOEDYKVTVALEGLDSSAAQVMAVYVNNLGYDGSGITIGIIDTGIDASHPDLOSKV | 180 | |
| Db | 121 | RFIOEDYKVTVALEGLDSSAAQVMAVYVNNLGYDGSGITIGIIDTGIDASHPDLOSKV | 180 | |
| QY | 181 | IGWVDFNNGRSYPRDDHGHTHVASIAAGTGAANGXYKGAAPAKIAGIVLADDSGS | 240 | |
| Db | 181 | IGWVDFNNGRSYPRDDHGHTHVASIAAGTGAANGXYKGAAPAKIAGIVLADDSGS | 240 | |
| QY | 241 | ISTTIKGVEMAVDNKDRTYGIKVTNLISGSSSDGTDSLQAVNNMADGIVVCVAAGNS | 300 | |

```

DB 241 ISTIIKGVMAVDNKKYGIKVINLSLSSQSSDGTDLASQAVNAMDAGLVVVAAGNS 300
QY 301 GGNITVGSPPAAASKVITVGAVDNNDIASFSSRGCTAGRLKPEVAVAGVDIIAPRAG 360
DB 301 GNNKTTIGSPAAASKVITVGAVDKDVITFSFSSRGCTAGRLKPEVAVAGVDIIAPRAG 360
QY 361 TSMGTPIINDYTTKASGTSMTATPHVSGVALILOAHPSWTPDKVKTALLETADIVAPKEIA 420
DB 361 TSMGPIINDYTTAAGTSMATPHVAGIALILOAHPSWTPDKVKTALLETADIVAPKEIA 420
QY 421 DIAYAGRVNVYKAIKYDDYAKLTFTGSVADKGSATHTPDSGATFTVATLYWDTGSSDI 480
DB 421 DIAYAGRVNVYKAIKYDDYAKLTFTGSVADKGSATHTPDSGATFTVATLYWDTGSSDI 480
QY 481 DLYIDPENGNDVSTAYYGFEEKYGYNPTAGTWTVKVYSGKGAANYOVVDSGLSQ 540
DB 481 DLYIDPENGNDVSTAYYGFEEKYGYNPTAGTWTVKVYSGKGAANYOVVDSGLSQ 540
QY 541 SGGGNPNPNPNPTTDTOTFTGTSVNDYMDTSDPTFMNNSGATKITGDLTFTDTSYND 600
DB 541 SGGGNPNPNPNPTTDTOTFTGTSVNDYMDTSDPTFMNNSGATKITGDLTFTDTSYND 600
QY 601 LDLYIDPENGNDVSTSSNSYEHVEYANPARGTWTFLVYASTYGMADYOLKAVVYQ 659
DB 601 LDLYIDPENGNDVSTSSNSYEHVEYANPARGTWTFLVYASTYGMADYOLKAVVYQ 659
QY 596 LDLYIDPENGNDVSTSSNSYEHVEYANPARGTWTFLVYASTYGMADYOLKAVVYQ 654

```

RESULT 7

US-09-841-553-1

Sequence 1, Application US/09841553

Publication No. US20020086402A1

GENERAL INFORMATION:

APPLICANT: TAKAKURA, Hikaru

MORISHITA, Mio

YAMAMOTO, Katsuniko

MITTA, Masanori

ASADA, Kiyozo

TSUNASAWA, Susumu

KATO, Ikunoshin

TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESS: Broadway and Newark

STREET: 419 Seventh Street N.W., Ste. 300

CITY: Washington

STATE: D.C.

COUNTRY: United States of America

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/841,553

FILING DATE: 24-Apr-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/894,818

FILING DATE: <Unknown>

APPLICATION NUMBER: JP 323285/1995

FILING DATE: 12-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Browdy, Roger L.

REGISTRATION NUMBER: 25,618

REFERENCE/DOCKET NUMBER: TAKAKURA=1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-5197

TELEFAX: (202) 737-3528

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 659 amino acids

TYPE: amino acid

```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-841-553-1

```

```

Query Match      85.0%; Score 2914; DB 3; Length 659;
Best Local Similarity 83.7%; Pred. No. 1.2e-197;
Matches 558; Conservative 47; Mismatches 46; Indels 16; Gaps 7;

```

```

QY 1 MKGLKALLIVLVGLVGVSAAPAEKVEQYRN---VEKNYGLTTPGLFRKIQKLNPRE 57
DB 1 MKRLGAVVLAIVLVELLGLTAAAPVKEV--VRNNAVOQKNKGLTTPGLFEKVQKRNMQ 58
QY 58 EISTVIFENHREKEIAVRVLELMGAKRYVYVHIIIPALADLKVBLLVLSGL--TG--G 113
DB 58 EVDIVIMFGSYGDRRAVAVLRLMGAQYKTSKIIIPAVAVKAKADLLIAGMIDTGYG 118
QY 114 KAKLSGVRFTQEDYKVTYSABLEGDESAAQVMATYVNN-LGYDGGITIGIITGIDAS 172
DB 119 NTRVSGIKFIQEDYKVVQDAA-----TSVSGIADPTVNNLSGYDGGVVAIVDTGIDAN 173
QY 173 HPDLGKVIYGVDFVNGSYPIYDGHGHTVASTIAGTGAASNGKTKGMAPAKAGIV 232
DB 174 HPDLGKVIYGVDAVNGSTPYDDQGHGHTVAGIYAGTGSV--NSQYIGVAPGAKVGVK 232
QY 223 LGADSGSISTTIKVENAVNDKKYGIKVINLSLSSQSSDGTSLGQAVNAMDAGIV 292
DB 223 LGADSGSVSTIIACVDVWQKDKYGIKVINLSLSSQSSDGTSLGQAVNAMDAGIV 292
QY 293 VCVAAAGNPGNTYTVGSPPAAASKVITVGAVDNNDIASFSSRGCTAGRLKPEVAVAGVD 352
DB 293 VCVAAAGNPGNTYTVGSPPAAASKVITVGAVDNNDIASFSSRGCTAGRLKPEVAVAGVD 352
QY 353 IIPRAGSGTSMGTPIINDYTTKASGTSMTATPHVSGVALILOAHPSWTPDKVKTALLETAD 412
DB 353 IIPRAGSGTSMGTPIINDYTTKASGTSMTATPHVSGVALILOAHPSWTPDKVKTALLETAD 412
QY 413 IIPRAGSGTSMGTPIINDYTTKASGTSMTATPHVSGVALILOAHPSWTPDKVKTALLETAD 472
DB 413 IIPRAGSGTSMGTPIINDYTTKASGTSMTATPHVSGVALILOAHPSWTPDKVKTALLETAD 472
QY 473 WDTGSSDIDLVLDPNGNEVDYSYAYYGFEEKYGYNPTAGTWTVKVSYKGAANYQVDV 532
DB 473 WDTGSSDIDLVLDPNGNEVDYSYAYYGFEEKYGYNPTAGTWTVKVSYKGAANYQVDV 532
QY 533 VSDGSLSGSGGNPNPNPNPTTDTOTFTGTSVNDYMDTSDPTFMNNSGATKITGDL 592
DB 533 VSDGSLSGSGGNPNPNPNPTTDTOTFTGTSVNDYMDTSDPTFMNNSGATKITGDL 592
QY 593 TFDTSYNDIDLVLDPNGNLVDRSTSSNSYEHVEYANPARGTWTFLVYASTYGMADYOL 652
DB 593 TFDTSYNDIDLVLDPNGNLVDRSTSSNSYEHVEYANPARGTWTFLVYASTYGMADYOL 652
QY 653 KAVVYVYQ 659
DB 653 KAVVYVYQ 659

```

RESULT 8

US-10-090-624-12

Sequence 12, Application US/10090624

Publication No. US2002013233A1

GENERAL INFORMATION:

APPLICANT: TAKAKURA, Hikaru

APPLICANT: MORISHITA, Mio

APPLICANT: SHIMOJO, Tomoko

APPLICANT: ASADA, Kiyozo

APPLICANT: KATO, Ikunoshin

TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE

FILE REFERENCE: TAKAKURA=6

CURRENT APPLICATION NUMBER: US/10/090,624

CURRENT FILING DATE: 2002-03-06

;; PRIOR APPLICATION NUMBER: 09/445,472
;; PRIOR FILING DATE: 1999-12-06
;; PRIOR APPLICATION NUMBER: 151969/1997
;; PRIOR FILING DATE: 1997-06-10
;; NUMBER OF SEQ ID NOS: 33
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 12
;; LENGTH: 659
;; TYPE: PR
;; ORGANISM: Thermococcus celer
US-10-090-624-12

Query Match 85.0%; Score 2914; DB 4; Length 659;
Best Local Similarity 83.7%; Pred. No. 1.2e-197;
Matches 558; Conservative 47; Mismatches 46; Indels 16; Gaps 7;

QY 1 MGLKALILVILVGLVGSVAAPBEKVEQVRN---VEKNYGLLTGGLFRKIQKLNPE 57
DB 1 MRLGAVVLAIVGLAGTALAAPVPRV--VRNNAVOQKNYGLLTGGLFRKVOQNNMNQ 58
QY 58 EISTVIVFENHREKEIAVRVLELMGAKRYVYHIIIPAIADLKVRDLVITSL--TG--G 113
DB 59 EVDVIMFGSYGRDRRAVKVRLMGAQVKSXKIIPAAVKIKARDLILAGMIDTGYFG 118
QY 114 KAKISGVRFIOEDYKVTVAELBGLDESAQVMATVYWN--LGYDGSGITIGIIDTGDAS 172
DB 119 NTRVSGIKRFIOEDYKVOVDA-----TSVSGIGADTVWNLSIGYDGSVVVAIVDTGIDAN 173
QY 173 HPDLGKVIQWVDPVNGRSYPYDDHGHGTHVASIAGTGAASNGKYKMAFGAKLAGIKY 232
DB 174 HPDLGKVIQWDAVNGRSTPYDDHGHGTHVAGIAGTGSV--NSQYIGVAPGAKLVGVKY 232
QY 233 LGADSGSISITIIKGVEMAVDNKDKYGIKYINLSLSSQSSDGTDSLQAVNNAMDAGIV 292
DB 233 LGADSGSVSTIIAGVDWVQNKDKYGIKYINLSLSSQSSDGTDSLQAVNNAMDAGIV 292
QY 293 VCVAAAGNSGPNITYVGSPPAASKVITVGAVDSDNNTASFSSRGPTADGRLKPEVAPGV 352
DB 293 VCVAAAGNSGPNITYVGSPPAASKVITVGAVDSDNNTASFSSRGPTADGRLKPEVAPGV 352
QY 353 IIPRAGSTMGTPINDYTTKASGTSMAATPHVSGVGLIILOAHPMSWTPDVKXTALIEFAD 412
DB 353 IIPRAGSTMGTPINDYTTKASGTSMAATPHVSGVGLIILOAHPMSWTPDVKXTALIEFAD 412
QY 413 IYAPKEIADIAYGAGRVNVYKAIKYDDYAKLFTGSVADKGSATHTFVSGATFTATLY 472
DB 413 IYAPKEIADIAYGAGRVNVYKAIKYDDYAKLFTGSVADKGSATHTFVSGATFTATLY 472
QY 473 WDTGSSDIDLXYLDPNGNEVDYSYTAAYGFEKVGYNPTAGTWTVKVSYKGAANYQVDV 532
DB 473 WDTGSSDIDLXYLDPNGNEVDYSYTAAYGFEKVGYNPTAGTWTVKVSYKGAANYQVDV 532
QY 533 VSDGSLSGSGGNPNPNPNPTPTDIOFTGSVNDYMDTSDPTFMNNSGATKITGDL 592
DB 533 VSDGSLSGSGGNPNPNPNPTPTDIOFTGSVNDYMDTSDPTFMNNSGATKITGDL 592
QY 593 TFDTSYNDLIDLXYLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLYASTYGMADYQL 652
DB 593 TFDTSYNDLIDLXYLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLYASTYGMADYQL 652
QY 653 KAVVYVG 659
DB 653 KAVVYVG 659

RESULT 9
US-10-800-684-1

;; Sequence 1, Application US/10800684
;; Publication No. US20050014221A1
;; GENERAL INFORMATION:
;; APPLICANT: TAKAKURA, Hikaru
;; APPLICANT: MORISHITA, Mio
;; APPLICANT: YAMAMOTO, Katsuhiko

;; APPLICANT: MITTA, Masanori
;; APPLICANT: ASADA, Kiyozo
;; APPLICANT: TSUNASAWA, Susumu
;; APPLICANT: KATO, Ikunoshin
;; TITLE OF INVENTION: ULTRATHERMOSTABLE PROTEASE GENES
;; FILE REFERENCE: TAKAKURA-1A
;; CURRENT APPLICATION NUMBER: US/10/800,684
;; PRIOR FILING DATE: 2004-03-16
;; PRIOR APPLICATION NUMBER: US/09/841,553
;; PRIOR FILING DATE: 2001-04-24
;; PRIOR APPLICATION NUMBER: 08/894,818
;; PRIOR FILING DATE: 1997-08-29
;; PRIOR APPLICATION NUMBER: JP2385/1995
;; PRIOR FILING DATE: 1995-12-12
;; PRIOR APPLICATION NUMBER: JP96/03253
;; PRIOR FILING DATE: 1996-11-07
;; NUMBER OF SEQ ID NOS: 45
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 1
;; LENGTH: 659
;; TYPE: PR
;; ORGANISM: Thermococcus celer
US-10-800-684-1

Query Match 85.0%; Score 2914; DB 5; Length 659;
Best Local Similarity 83.7%; Pred. No. 1.2e-197;
Matches 558; Conservative 47; Mismatches 46; Indels 16; Gaps 7;

QY 1 MGLKALILVILVGLVGSVAAPBEKVEQVRN---VEKNYGLLTGGLFRKIQKLNPE 57
DB 1 MRLGAVVLAIVGLAGTALAAPVPRV--VRNNAVOQKNYGLLTGGLFRKVOQNNMNQ 58
QY 58 EISTVIVFENHREKEIAVRVLELMGAKRYVYHIIIPAIADLKVRDLVITSL--TG--G 113
DB 59 EVDVIMFGSYGRDRRAVKVRLMGAQVKSXKIIPAAVKIKARDLILAGMIDTGYFG 118
QY 114 KAKISGVRFIOEDYKVTVAELBGLDESAQVMATVYWN--LGYDGSGITIGIIDTGDAS 172
DB 119 NTRVSGIKRFIOEDYKVOVDA-----TSVSGIGADTVWNLSIGYDGSVVVAIVDTGIDAN 173
QY 173 HPDLGKVIQWVDPVNGRSYPYDDHGHGTHVASIAGTGAASNGKYKMAFGAKLAGIKY 232
DB 174 HPDLGKVIQWDAVNGRSTPYDDHGHGTHVAGIAGTGSV--NSQYIGVAPGAKLVGVKY 232
QY 233 LGADSGSISITIIKGVEMAVDNKDKYGIKYINLSLSSQSSDGTDSLQAVNNAMDAGIV 292
DB 233 LGADSGSVSTIIAGVDWVQNKDKYGIKYINLSLSSQSSDGTDSLQAVNNAMDAGIV 292
QY 293 VCVAAAGNSGPNITYVGSPPAASKVITVGAVDSDNNTASFSSRGPTADGRLKPEVAPGV 352
DB 293 VCVAAAGNSGPNITYVGSPPAASKVITVGAVDSDNNTASFSSRGPTADGRLKPEVAPGV 352
QY 353 IIPRAGSTMGTPINDYTTKASGTSMAATPHVSGVGLIILOAHPMSWTPDVKXTALIEFAD 412
DB 353 IIPRAGSTMGTPINDYTTKASGTSMAATPHVSGVGLIILOAHPMSWTPDVKXTALIEFAD 412
QY 413 IYAPKEIADIAYGAGRVNVYKAIKYDDYAKLFTGSVADKGSATHTFVSGATFTATLY 472
DB 413 IYAPKEIADIAYGAGRVNVYKAIKYDDYAKLFTGSVADKGSATHTFVSGATFTATLY 472
QY 473 WDTGSSDIDLXYLDPNGNEVDYSYTAAYGFEKVGYNPTAGTWTVKVSYKGAANYQVDV 532
DB 473 WDTGSSDIDLXYLDPNGNEVDYSYTAAYGFEKVGYNPTAGTWTVKVSYKGAANYQVDV 532
QY 533 VSDGSLSGSGGNPNPNPNPTPTDIOFTGSVNDYMDTSDPTFMNNSGATKITGDL 592
DB 533 VSDGSLSGSGGNPNPNPNPTPTDIOFTGSVNDYMDTSDPTFMNNSGATKITGDL 592
QY 593 TFDTSYNDLIDLXYLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLYASTYGMADYQL 652
DB 593 TFDTSYNDLIDLXYLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLYASTYGMADYQL 652
QY 653 KAVVYVG 659
DB 653 KAVVYVG 659

Db 653 KAVVYVG 659

RESULT 10
US-10-888-588-12
Sequence 12, Application US/10888588
Publication No. US20050084934A1
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOTO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/10/888,588
CURRENT FILING DATE: 2004-07-12
PRIOR FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: US/10/090,624
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 09/445,472
PRIOR FILING DATE: 1997-06-10
PRIOR APPLICATION NUMBER: 151969/1997
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 12
LENGTH: 659
TYPE: PRT
ORGANISM: Thermococcus celer
US-10-888-588-12

Query Match 85.0%; Score 2914; DB 5; Length 659;
Best Local Similarity 83.7%; Pred. No. 1,2e-197;
Matches 558; Conservative 47; Mismatches 46; Indels 16; Gaps 7;

1 MKGLKALIVILVGLGVSVAAAPKKEVOYRN--VEKNYGLTPGLFRKIQLNPNE 57
1 MRLGAVNALVGLVGLAGTALAAPYKPV--VRNNAVQKNGYGLTPGLFKVYQVRNNMG 58
58 EISTVIFENHREKEIAVLEIMGAKVRYVYHIIIPAIADLKVRDLVLSGL--TG--G 113
59 EVDTVIMFGSYGRDRAVKVLRMGQVQKSYKIIIPAVAVKIKARDLLIAGMIDGYRG 118
114 KAKLSVRIQEDRYKTVSAELEGDESAQVMATVYV--LGYDGGITIGIIDGIDS 172
119 NTRVSGIKRFOEDYKQVDA-----TSVQIGADVWNSLGYDGGVVAIVDTGIDAN 173
173 HPDLQKVIQWVDFVNGRSYPYDDHGHVHVASIAAGTGAASNGYKGMAPGAKLAGIV 232
174 HPDLKQKVIQWDAVNGRSTPYDDHGHVHAGIVAGTSV--NSQITGVAFGAKLVGAV 232
233 LGADSGSISTTIKGYEAVDNKDKYIKVINSLSGSSQSSDGTSLSGAVNNAMADGIV 292
233 LGADSGSVSTIIAGVDMVQNKDKYIRVINSLSGSSQSSDGTSLSGAVNNAMADGIV 292
293 VCVAAAGSGPNTYVTSPPAAASKVITVGVAVDSNDNIASFSSRPTADGRLKEVVAAGVD 352
293 VCVAAAGSGPNTYVTSPPAAASKVITVGVAVDSNDNIASFSSRPTADGRLKEVVAAGVD 352
353 IIAPRASGTSMTPIINDYTTKASGTSMTATPHVSGVALLLOAHPSHTPPKXTALLETAD 412
413 IIVAPKEIADIAVAGAGRVNYKAIKYDVAKLFTGSGVADKSGATHTFDVSGATFVATLY 472
413 IIVAPKEIADIAVAGAGRVNYKAIKYDVAKLFTGSGVADKSGATHTFDVSGATFVATLY 472
473 WDTGSSDIDLILYDPNGNEVDYSYTAIYGEPEKYGYNPTAGTIVVYVSGKAAANYQVY 532
473 WDTGSSDIDLILYDPNGNEVDYSYTAIYGEPEKYGYNPTAGTIVVYVSGKAAANYQVY 532
533 VSDGSLSGGCGGNPNPNPPTTTDTGTGTSVNDYDTSPTFTMANVNSGATKITGDL 592

Db 533 VSDGSLSGGCGGNPNPNPPTTTDTGTGTSVNDYDTSPTFTMANVNSGATKITGDL 592

593 TFDTSYNLDLILYDPNGNLVDRSTSSNSYEHVEYANPAPGTWTEFLVYASTYGNADYOL 652
593 TFDTSYNLDLILYDPNGNLVDRSTSSNSYEHVEYANPAPGTWTEFLVYASTYGNADYOL 652

Query 653 KAVVYVG 659

RESULT 11
US-09-841-553-3
Sequence 3, Application US/09841553
Publication No. US20020086402A1
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MITTA, Masanori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
KATO, Ikunoshin
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Broadway and Neilmark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/841,553
FILING DATE: 24-Apr-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/894,818
FILING DATE: <Unknown>
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-Dec-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA-1
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 522 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION:
Xaa at position 428 is Gly or Val.
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-841-553-3

Query Match 68.5%; Score 2349.5; DB 3; Length 522;
Best Local Similarity 85.2%; Pred. No. 9e-158;
Matches 449; Conservative 22; Mismatches 51; Indels 5; Gaps 1;

1 AELGELDESAQAQVMATYVWNLGYDGGITIGIIDGIDASHPDLQKVIQWVDFVNGRSY 192
1 AELGELDESAQAQVMATYVWNLGYDGGITIGIIDGIDASHPDLQKVIQWVDFVNGRSY 60

```
QY 193 PYDDHGHGTHVASIAAGTGAASNGKYKGMAPGAKLAGIKVLGADSGSISTIIKGVEMAV 252
Db 61 PYDDHGHGTHVASIAAGTGAASNGKYKGMAPGAKLAGIKVLGADSGSISTIIKGVEMAV 120
QY 253 DNKDKYGIKVINLSIGSSQSDGTDSLGAOVNANMADGIYVCVAAGNSGPNYTVGSPAA 312
Db 121 DNKDKYGIKVINLSIGSSQSDGTDSLGAOVNANMADGIYVCVAAGNSGPNYTVGSPAA 180
QY 313 ASKVIITGAVDNDNINASFSSRGPTADGRLEKEVVAAPGVDIAPRASGTSMTGPIINDYTT 372
Db 181 ASKVIITGAVDNDNINASFSSRGPTADGRLEKEVVAAPGVDIAPRASGTSMTGPIINDYTT 240
QY 373 KASGTSMAATPHVSGVGAIILOAHPSWTPDKYKTALIEADIAVPAKEIADIAVAGAVVAVY 432
Db 241 AAPGTSMAATPHVAGIALLLOAHPSWTPDKYKTALIEADIAVPAKEIADIAVAGAVVAVY 300
QY 433 KAIKYDDYAKLFTFGSVADKSGATHTFDVSGATFTVATLYMDTSSDIDLXYDPNGNEV 492
Db 301 KAIKYDDYAKLFTFGSVADKSGATHTFDVSGATFTVATLYMDTSSDIDLXYDPNGNEV 360
QY 493 DYSYTAAYGFEKVGYYNPAGTWTWKVSYKGAANYOVVSDGSLSGSGGNPNPNP 552
Db 361 DYSYTAAYGFEKVGYYNPAGTWTWKVSYKGAANYOVVSDGSLSGSGGNPNPNP 415
QY 553 NPPTTDTGTFTGSDVNDYMDTSDFTTMVNSGATKITGDLTFDTSYNDLXYDPNGNL 612
Db 416 QPEPTVDAKTFQXSDHYHYDSDFTFTMTVNSGATKITGDLTFDTSYNDLXYDPNGNL 475
QY 613 VDRSTSSNSYEHVEYANAPAGTWTFLVAYASTYGMAVYQLAKAVVYVG 659
Db 476 VDRSESPNSYEHVEYLTLPAGTWTFLVAYASTYGMAVYELTAKVYVG 522
```

```
RESULT 12
US-10-090-624-4
; Sequence 4, Application US/10090624
; Publication No. US2002012335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (428)..(428)
; OTHER INFORMATION: Xaa at position 428 is Gly or Val.
US-10-090-624-4
```

```
Query Match 68.5%; Score 2349.5; DB 4; Length 522;
Best Local Similarity 85.2%; Pred. No. 9e-158;
Matches 449; Conservative 22; Mismatches 51; Indels 5; Gaps 1;
QY 133 AELBGLDESAQAQVMATYVMNLGYDGSIGITIGIITDGDASHPDLOGKVIQWDFVNGRSY 192
Db 1 AELBGLDESAQAQVMATYVMNLGYDGSIGITIGIITDGDASHPDLOGKVIQWDFVNGRSY 60
QY 193 PYDDHGHGTHVASIAAGTGAASNGKYKGMAPGAKLAGIKVLGADSGSISTIIKGVEMAV 252
```

```
Db 61 PYDDHGHGTHVASIAAGTGAASNGKYKGMAPGAKLAGIKVLGADSGSISTIIKGVEMAV 120
QY 253 DNKDKYGIKVINLSIGSSQSDGTDSLGAOVNANMADGIYVCVAAGNSGPNYTVGSPAA 312
Db 121 DNKDKYGIKVINLSIGSSQSDGTDSLGAOVNANMADGIYVCVAAGNSGPNYTVGSPAA 180
QY 313 ASKVIITGAVDNDNINASFSSRGPTADGRLEKEVVAAPGVDIAPRASGTSMTGPIINDYTT 372
Db 181 ASKVIITGAVDNDNINASFSSRGPTADGRLEKEVVAAPGVDIAPRASGTSMTGPIINDYTT 240
QY 373 KASGTSMAATPHVSGVGAIILOAHPSWTPDKYKTALIEADIAVPAKEIADIAVAGAVVAVY 432
Db 241 AAPGTSMAATPHVAGIALLLOAHPSWTPDKYKTALIEADIAVPAKEIADIAVAGAVVAVY 300
QY 433 KAIKYDDYAKLFTFGSVADKSGATHTFDVSGATFTVATLYMDTSSDIDLXYDPNGNEV 492
Db 301 KAIKYDDYAKLFTFGSVADKSGATHTFDVSGATFTVATLYMDTSSDIDLXYDPNGNEV 360
QY 493 DYSYTAAYGFEKVGYYNPAGTWTWKVSYKGAANYOVVSDGSLSGSGGNPNPNP 552
Db 361 DYSYTAAYGFEKVGYYNPAGTWTWKVSYKGAANYOVVSDGSLSGSGGNPNPNP 415
QY 553 NPPTTDTGTFTGSDVNDYMDTSDFTTMVNSGATKITGDLTFDTSYNDLXYDPNGNL 612
Db 416 QPEPTVDAKTFQXSDHYHYDSDFTFTMTVNSGATKITGDLTFDTSYNDLXYDPNGNL 475
QY 613 VDRSTSSNSYEHVEYANAPAGTWTFLVAYASTYGMAVYQLAKAVVYVG 659
Db 476 VDRSESPNSYEHVEYLTLPAGTWTFLVAYASTYGMAVYELTAKVYVG 522
```

```
RESULT 13
US-10-800-684-3
; Sequence 3, Application US/10800684
; Publication No. US2005001422A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MITTA, Masanori
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: ULTRATHERMOSTABLE PROTEASE GENES
; FILE REFERENCE: TAKAKURA-1A
; CURRENT APPLICATION NUMBER: US/10/800,684
; PRIOR FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: US/09/841,553
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 08/894,818
; PRIOR FILING DATE: 1997-08-29
; PRIOR APPLICATION NUMBER: JP32385/1995
; PRIOR FILING DATE: 1995-12-12
; PRIOR APPLICATION NUMBER: JP96/03253
; PRIOR FILING DATE: 1996-11-07
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (428)..(428)
; OTHER INFORMATION: Xaa is Gly or Val
US-10-800-684-3
```

```
Query Match 68.5%; Score 2349.5; DB 5; Length 522;
Best Local Similarity 85.2%; Pred. No. 9e-158;
Matches 449; Conservative 22; Mismatches 51; Indels 5; Gaps 1;
QY 133 AELBGLDESAQAQVMATYVMNLGYDGSIGITIGIITDGDASHPDLOGKVIQWDFVNGRSY 192
```

```
Db 1 ALEGLDESAQAQVMAVYWNLYGDSGIIITGIDGIDASHPDLOGKVIQWDFVNGRSY 60
QY 193 PYDDHGHGTHVASIAAGTGAASNGKYKGMAPAKLAGIKVILGADSGSISTIIKVENAV 252
Db 61 PYDDHGHGTHVASIAAGTGAASNGKYKGMAPAKLAGIKVILGADSGSISTIIKVENAV 120
QY 253 DNKDKYGIKVINLSIGSSQSSDGTDLISQAVNANMADAGIYVCAAGNSGPNITYVGPAA 312
Db 121 DNKDKYGIKVINLSIGSSQSSDGTDLISQAVNANMADAGIYVCAAGNSGPNITYVGPAA 180
QY 313 ASKVTIVGAVDNSNDNIASFSSRGPTADGRLKPEVVAPEGVDIIAPRAGSTSGTPIINDYTT 372
Db 181 ASKVTIVGAVDNSNDNIASFSSRGPTADGRLKPEVVAPEGVDIIAPRAGSTSGTPIINDYTT 240
QY 373 KASGSMATPHVSGVALLILOAHPSMTDPKXTALIEADIAPEIADIAAGRVNAY 432
Db 241 AAPGSMATPHVAGIAALLILOAHPSMTDPKXTALIEADIAPEIADIAAGRVNAY 300
QY 433 KAIKDYDVAKLFTGSGVADKSGATHTFDVSGATFVTATLYMDTGSSDIDLXYLDPNNGEV 492
Db 301 KAINDNVAKLFTGSGVADKSGATHTFDVSGATFVTATLYMDTGSSDIDLXYLDPNNGEV 360
QY 493 DYSYTAAYGFEKVGYYNPTAGTWTYKVSYKGAANYQVDVSDGSLSGSGGPNPNPNP 552
Db 361 DYSYTAAYGFEKVGYYNPTAGTWTYKVSYKGAANYQVDVSDGSLSGSGGPNPNPNP 415
QY 553 NPTPTDQOTFPGSVNDVSDTFTTMVNSGATKITGDLTFTDSYNDLXYLDPNNGVL 612
Db 416 QPEPTVADKTFQXSDHYIYDRSDTFTTMVNSGATKITGDLTFTDSYNDLXYLDPNNGVL 475
QY 613 VDRSTSSNHYEHEVYANPAGTWTFLVYAYSTYGADYOLKAVVYVG 659
Db 476 VDRSESPNSYHEVYLTAPAGTWTFLVYAYSTYGADYOLKAVVYVG 522
```

```
RESULT 14
US-10-888-588-4
; Sequence 4, Application US/10888588
; Publication No. US20050084934A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hi-karu
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOTO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/888,588
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: (428)..(428)
; OTHER INFORMATION: Xaa at position 428 is Gly or Val.
US-10-888-588-4
```

```
Query Match 68.5%; Score 2349.5; DB 5; Length 522;
Best Local Similarity 85.2%; Pred. No. 9e-158;
Matches 449; Conservative 22; Mismatches 51; Indels 5; Gaps 1;
QY 133 ALEGLDESAQAQVMAVYWNLYGDSGIIITGIDGIDASHPDLOGKVIQWDFVNGRSY 192
```

```
Db 1 ALEGLDESAQAQVMAVYWNLYGDSGIIITGIDGIDASHPDLOGKVIQWDFVNGRSY 60
QY 193 PYDDHGHGTHVASIAAGTGAASNGKYKGMAPAKLAGIKVILGADSGSISTIIKVENAV 252
Db 61 PYDDHGHGTHVASIAAGTGAASNGKYKGMAPAKLAGIKVILGADSGSISTIIKVENAV 120
QY 253 DNKDKYGIKVINLSIGSSQSSDGTDLISQAVNANMADAGIYVCAAGNSGPNITYVGPAA 312
Db 121 DNKDKYGIKVINLSIGSSQSSDGTDLISQAVNANMADAGIYVCAAGNSGPNITYVGPAA 180
QY 313 ASKVTIVGAVDNSNDNIASFSSRGPTADGRLKPEVVAPEGVDIIAPRAGSTSGTPIINDYTT 372
Db 181 ASKVTIVGAVDNSNDNIASFSSRGPTADGRLKPEVVAPEGVDIIAPRAGSTSGTPIINDYTT 240
QY 373 KASGSMATPHVSGVALLILOAHPSMTDPKXTALIEADIAPEIADIAAGRVNAY 432
Db 241 AAPGSMATPHVAGIAALLILOAHPSMTDPKXTALIEADIAPEIADIAAGRVNAY 300
QY 433 KAIKDYDVAKLFTGSGVADKSGATHTFDVSGATFVTATLYMDTGSSDIDLXYLDPNNGEV 492
Db 301 KAINDNVAKLFTGSGVADKSGATHTFDVSGATFVTATLYMDTGSSDIDLXYLDPNNGEV 360
QY 493 DYSYTAAYGFEKVGYYNPTAGTWTYKVSYKGAANYQVDVSDGSLSGSGGPNPNPNP 552
Db 361 DYSYTAAYGFEKVGYYNPTAGTWTYKVSYKGAANYQVDVSDGSLSGSGGPNPNPNP 415
QY 553 NPTPTDQOTFPGSVNDVSDTFTTMVNSGATKITGDLTFTDSYNDLXYLDPNNGVL 612
Db 416 QPEPTVADKTFQXSDHYIYDRSDTFTTMVNSGATKITGDLTFTDSYNDLXYLDPNNGVL 475
QY 613 VDRSTSSNHYEHEVYANPAGTWTFLVYAYSTYGADYOLKAVVYVG 659
Db 476 VDRSESPNSYHEVYLTAPAGTWTFLVYAYSTYGADYOLKAVVYVG 522
```

```
RESULT 15
US-10-090-624-1
; Sequence 1, Application US/10090624
; Publication No. US2002013235A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hi-karu
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOTO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-10-090-624-1
```

```
Query Match 56.0%; Score 1918; DB 4; Length 412;
Best Local Similarity 89.5%; Pred. No. 2.4e-127;
Matches 367; Conservative 15; Mismatches 28; Indels 0; Gaps 0;
QY 133 ALEGLDESAQAQVMAVYWNLYGDSGIIITGIDGIDASHPDLOGKVIQWDFVNGRSY 192
Db 1 ALEGLDESAQAQVMAVYWNLYGDSGIIITGIDGIDASHPDLOGKVIQWDFVNGRSY 60
QY 193 PYDDHGHGTHVASIAAGTGAASNGKYKGMAPAKLAGIKVILGADSGSISTIIKVENAV 252
Db 61 PYDDHGHGTHVASIAAGTGAASNGKYKGMAPAKLAGIKVILGADSGSISTIIKVENAV 120
```


QY 253 DNRKXGIXVNLISLSSQSSDGTSLQA VNNAMDAGIVVYAAGNSGPNITYVGSPPA 312
| | | | | : | | | | | : | | | | | : | | | | |
Db 121 DNRKXGIXVNLISLSSQSSDGTSLQA VNNAMDAGLVVVAAGNSGPNKTYTIGSPAA 180
| | | | | : | | | | | : | | | | | : | | | | |
QY 313 ASKVIIVGAVDSNDNTIASFSSRGPTADGRLEPEVVAFCVDIIPRASGTSMTPIINDYTT 372
| | | | | : | | | | | : | | | | | : | | | | |
Db 181 ASKVIIVGAVDKYDVITSFSSRGPTADGRLEPEVVAFCVNIIPRASGTSMTGQPIINDYTT 240
| | | | | : | | | | | : | | | | | : | | | | |
QY 373 KASGTSMTPEHVSQVGLIIOAHPSWTPDKVKTALITETADIIVAPKEIADIAYGAGRNVY 432
| | | | | : | | | | | : | | | | | : | | | | |
Db 241 AAPGTSMTPEHVAIGIALLIQAHPSWTPDKVKTALITETADIIVAPKEIADIAYGAGRNVY 300
| | | | | : | | | | | : | | | | | : | | | | |
QY 433 KAIKYDDYAKLTFTGSVADKGSATHTFVSGATFVTATLYMDTSSDPDLXYDPNGNEV 492
| | | | | : | | | | | : | | | | | : | | | | |
Db 301 KAINYDNYAKLVFTGYVANKSQTHQFVISGASFVTATLYMDNANSDDLXYDPNGNCV 360
| | | | | : | | | | | : | | | | | : | | | | |
QY 493 DYSYTAIYGFGEKGYNPTAGTWTVKVVSYKGAANTQVDVVSDDSLSQSG 542
| | | | | : | | | | | : | | | | | : | | | | |
Db 361 DYSYTAIYGFGEKGYNPTAGTWTIKVVSYSGSANTQVDVVSDDSLSQSG 410
| | | | | : | | | | | : | | | | | : | | | | |

Search completed: January 6, 2007, 22:37:51
Job time : 125.5 secs


```

NAME/KEY: misc.feature
LOCATION: (1128499)..(1128499)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1128505)..(1128506)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1128517)..(1128518)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1128539)..(1128540)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1561400)..(1561400)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1561477)..(1561477)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1561545)..(1561545)
OTHER INFORMATION: n is a or c or g or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1767941)..(1767941)
OTHER INFORMATION: n is a or c or g or t.
US-10-526-324-251

```

```

Query Match 78.9%; Score 2705; DB 6; Length 663;
Best Local Similarity 77.7%; Pred. No. 1.4e-182;
Matches 517; Conservative 60; Mismatches 80; Indels 8; Gaps 4;

```

```

QY 1 MKGKALIVLVGIVGVSAAAPKKVDEQYRV--EKYGLTTPGLFKIQKLPNEE 58
DB 1 MKKGAVALVLFVLGVLGSLAAPKPA--VRNVSOQKNGYGLTTPGLFKVQRMWDCE 58
QY 59 ISTVIVENHREKEIAVRLVLMGAKRYVYHIIIPALADLKYRDLVLSGL--TG--GK 114
DB 59 VSTIMFDNQADKKAKEIIDLFLGAKIKNYHIIIPALAVIKYKDLIIIGLMDTGYFGN 118
QY 115 AKLSGVRFIQEDYKVTYSAELBGLDESAQVMATYNNIGYDGGITIGIIDGIDASHP 174
DB 119 AQLSGVQFIQEDYVYKVAVETEGIDBSAAQVMATNMNIGYDGGITIGIIDGIDASHP 178
QY 175 DLQOKVIGWDFVNGRGPYDDHGHGTHVASIAGTGAASNGYKGMAPAKLAGIKVVG 234
DB 179 DLQOKVIGWDFVNGKTTPTDDHGHGTHVASIAGTGAASNGYKGMAPAKLAGIKVVG 238
QY 235 ADGSGSISTIIKGVEMAVDNKDKYKIKVINLSIGSSQSSDGTDSLQAVNANMADAGIYVC 294
DB 239 GQSGSISDIINGVDMAVNQKDKYKIKVINLSIGSSQSSDGTDSLQAVNANMADAGIYVC 298
QY 295 VAANGSGPNTYTGSPPAASKVITVGAVDSNDNIASFSSRGPADGLKPEVVAPEVDII 354
DB 299 VAANGSGPNTYTGSPPAASKVITVGAVDYITTFSSRGPADNRLKPEVVAPEVVAPEVDII 358
QY 355 APRAAGTSMGTPINDYTTKASGTSMATPHYSGVGLIIOAHPEWPDOKYKTLIETRADIV 414
DB 359 AARASGTSMGTPINDYTTAAPTSMATPHYAGIALLQAHPEWPDOKYKTLIETRADIV 418
QY 415 APPEIADIAGAGRVNYYKAIKYDYAKLFTGVSADKGSATTFPVSAGTEVATLYMD 474
DB 419 KPPEIADIAGAGRVNYYKAAYDYAKLFTGYVSNKGSQSHQFTISAGFVATLYMD 478
QY 475 TGSSDIDLIVYDNGEVDYSYTAYYGFKEVNGYNTTACTGTAKYKVSYKGAANYQYDVVS 534
DB 479 NGSGDIDLIVYDNGEVDYSYTAYYGFKEVNGYNTTACTGTAKYKVSYKGAANYQYDVVS 538

```

```

QY 535 DGSLSQSGGNPNPNPPTTPTDQTFGVSNDYWDSDTFETMNVNSGATKITGDLTF 594
DB 539 DGSLSQSGGSGGSPSPSPVDEKTFGTGVHDYDSDTFTMNVNSGATKITGDLTF 598
QY 595 DTSYNDIDLIVDPNGNLVDRSTSSNSYEHVEYANPAPGTWELVYAYSTYGNADYQLKA 654
DB 599 DTSYNDIDLIVDPNGNLVDRSSSSNSYEHVEYANNPAPGTWELVYAYDYGYADYQLDA 658
QY 655 VVYVG 659
DB 659 KVIYG 663

```

RESULT 2

```

US-10-530-643-8
Sequence 8, Application US/10530643
Publication No. US2006025995A1
GENERAL INFORMATION:
APPLICANT: CAYOURETTE, Michelle
APPLICANT: HANSEN, Connie Jo
APPLICANT: MCCLEURE, Amy
APPLICANT: SUN, May
APPLICANT: GRAMATIKOVA, Svetlana
APPLICANT: DYCAICO, Mark
APPLICANT: BARTON, Nelson R.
APPLICANT: STEGE, Justin T.
APPLICANT: ABOUSHADI, Nahla M.
TITLE OF INVENTION: PROTEASES, NUCLEIC ACIDS ENCODING THEM
FILE REFERENCE: 564462004100
CURRENT APPLICATION NUMBER: US/10/530,643
PRIOR FILING DATE: 2006-04-07
PRIOR APPLICATION NUMBER: PCT/US03/32819
PRIOR FILING DATE: 2003-10-10
PRIOR APPLICATION NUMBER: 60/471,423
PRIOR FILING DATE: 2003-05-16
PRIOR APPLICATION NUMBER: 60/418,467
PRIOR FILING DATE: 2002-10-10
NUMBER OF SEQ ID NOS: 255
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 443
TYPE: PRT
ORGANISM: Bacteria
US-10-530-643-8

```

```

Query Match 19.5%; Score 670; DB 6; Length 443;
Best Local Similarity 37.2%; Pred. No. 2.5e-39;
Matches 161; Conservative 81; Mismatches 149; Indels 42; Gaps 11;

```

```

QY 30 EGVNRVENKYGILTPGLFKIQKLPN--EISTVIVENHREKEIAVRLVLM----GAK 84
DB 23 DRLNLVYKPFKMTPEFLKMFEPFKRKRKLVILEFSSNCQEGCCLVENQAFTGRNCK 82
QY 85 VAYVYHIIIPALADLKYRDLVLSGLTGKAKLSCVRFIQEDYKVTYSAELBGLDESA-- 142
DB 83 IKNEFSIISCSADITPSEVL-----BEVLTWCNHIK--KYLLNHEVRALLDTAVI 130
QY 143 -AQWATYNNI.GYGVSGITIGIIDGIDASHGPDLOGVIGWDFVNGRSPYDDHGHGT 201
DB 131 SANAKIYRANTYLTGKITTAVITDGL--FPHITLSGIIIFVDFINRITETTDNGHGT 189
QY 202 HVASIAAGTGAASNGYKGMAPGAKLAGIKYLGADGSGSISTIIKGVEMAV---DNKDKY 258
DB 190 HCAQDALDGDASGGLYMGAPAEANVIGVYKLVNKVGSGLTETWQGVDMCIRYDQNDQ 249
QY 259 GIKVINLSIGS---SQSSDGTDSLQAVNANMADAGIYVCVAAGSGPNTYTGSPPAASK 315
DB 250 KINIINSLGAPAPRYENENDPVMYKVEKAMENGIYVCVAAGSGEPASTIASPVSSEQ 309
QY 316 VITVGAVD-----SNDNIASFSSRGPADGLKPEVVAPEVDIAPRASGTM----- 363
DB 310 VITVGALDDKTTADTRSDDEVAFSRSKQPTIYKVKYDIIAPGVDTISLSPNSYLDKQ 369

```

```

Qy 364 -GTPINDYTKASGTMATPHVSGVALILOAHPSWTPTALLETADIVAPKEIADI 422
Db 370 KGNRGSDYFSLSGTSMATPICAGIALILOHNPATPQEVLDLKRGTDLMTNRD--PN 427
Qy 423 AYAGRVNYKAI 435
Db 428 IYAGYINAMNSV 440

RESULT 3
US-10-530-643-6
; Sequence 6, Application US/10530643
; Publication No. US2006025995A1
; GENERAL INFORMATION:
; APPLICANT: CAYOUEITE, Michelle
; APPLICANT: HANSEN, Connie Jo
; APPLICANT: MCCURE, Amy
; APPLICANT: SUN, May
; APPLICANT: GRAMATIKOVA, Svetlana
; APPLICANT: DYCAICO, Mark
; APPLICANT: BARTON, Nelson R.
; APPLICANT: STEGE, Justin T.
; APPLICANT: ABOUSHADI, Nahla M.
; TITLE OF INVENTION: PROTEASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 564462004100
; CURRENT APPLICATION NUMBER: US/10/530,643
; PRIOR FILING DATE: 2006-04-07
; PRIOR APPLICATION NUMBER: PCT/US03/32819
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/471,423
; PRIOR FILING DATE: 2003-05-16
; PRIOR APPLICATION NUMBER: 60/419,467
; NUMBER OF SEQ ID NOS: 255
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1227
; TYPE: PRT
; ORGANISM: Bacteria
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(32)
US-10-530-643-6

Query Match 19.3%; Score 660; DB 6; Length 1227;
Best Local Similarity 45.1%; Pred. No. 5,4e-38;
Matches 157; Conservative .52; Mismatches 115; Indels 24; Gaps 11;

Qy 140 ESAOVMATYVNNLGYDGGITIGIITGIDASHPDLOGKVGWDFVNGRSPYDDHGH 199
Db 201 ESNAGIGPKAWAGITGKGVAVAVLDTGADLAHPDLAGRVTESKSFINGQVVA-DRNCH 259
Qy 200 GTHVVASIAGTGAASNGKYGMAPGAAGKLAGIKVILGADSGSISTIIKGVEMVNDKDXG 259
Db 260 GTHVASTYGGSGSGGKEKGVAPGATILAVKVLSDGSGSEKIIAGEMAAKIDID-- 316
Qy 260 IKVINLSGSSOSDGTDSLQAVNN-AMDAGIWCVAAAGNSG-PWTYVGSPPAAASKYI 317
Db 317 AKIVMSIGSRSPDGTDMALAVNTLLETGALFVIAAGNSGYG--SIGSPGAADSL 374
Qy 318 TYGAVDSNDNINASFSGRPT-ADGRLKEPVVAPGVDIAPRASGYSMGTPINDYTKASG 376
Db 375 TIGAVDSADDEAAVFTSOGPRYGDOLKPLDLSAPVDIILAR-----SOLIPGSLYTSMSG 430
Qy 377 TSMATPHVSGVALILOAHPSWTPTALLETADIVAPKEIADIYAGRVNYKAIK 436
Db 431 TSMATPHVAGVALLAERHPDWTGAOLKDALMSSSKTL--DASSALDSGRVDAALAA 487
Qy 437 YDDYAKLTPTGSAVDKGSATHTFDVSGATFTVATILYMDTGSSDIDLYL 484
Db 488 ----ANVTATGS-ADLGFVAMPYASKPVTKIVTY---TNSSDAPVEL 527

```

```

RESULT 4
US-10-530-643-124
; Sequence 124, Application US/10530643
; Publication No. US2006025995A1
; GENERAL INFORMATION:
; APPLICANT: CAYOUEITE, Michelle
; APPLICANT: HANSEN, Connie Jo
; APPLICANT: MCCURE, Amy
; APPLICANT: SUN, May
; APPLICANT: GRAMATIKOVA, Svetlana
; APPLICANT: DYCAICO, Mark
; APPLICANT: BARTON, Nelson R.
; APPLICANT: STEGE, Justin T.
; APPLICANT: ABOUSHADI, Nahla M.
; TITLE OF INVENTION: PROTEASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 564462004100
; CURRENT APPLICATION NUMBER: US/10/530,643
; PRIOR FILING DATE: 2006-04-07
; PRIOR APPLICATION NUMBER: PCT/US03/32819
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/471,423
; PRIOR FILING DATE: 2003-05-16
; PRIOR APPLICATION NUMBER: 60/419,467
; NUMBER OF SEQ ID NOS: 255
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 124
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample.
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(56)
US-10-530-643-124

Query Match 17.2%; Score 588; DB 6; Length 552;
Best Local Similarity 39.1%; Pred. No. 2,1e-33;
Matches 155; Conservative .46; Mismatches 106; Indels 89; Gaps 12;

Qy 87 YYHIIIPAIADLKRDL--LVISGLTGKAKLGSVRPIQIEYKTVASALEGLDEBSAAQ 144
Db 198 YKVNHPRRRVSIALQDVPPIODGIIAGRVDEG-----NDCELTKRCL----- 243
Qy 145 WATVVMNLGYDGGITIGIITGIDASHPDLOG-----KVLGWDFVNGRSPY 192
Db 244 -----TGEVYTAIIDTGVDTYHPDLGCTTGQEFGLGACEKVIYGVDFINDD 292
Qy 193 PYDHGHTHVASIAGTGAASNGKYGMAPGAAGKLAGIKVILGADSGSISTIIKGVEMV 252
Db 223 AMDMDGHGHVAGIAAG-----NGLKGVAPDAKILAVKVLGAGGYGTGEGIIAGIEQAV 347
Qy 223 DNKDKYGIKVINLSGSSOSDGTDSL-QAVNNAMDAGIWCVAAAGNSGPTTYVGSAT 311
Db 348 ID---GADIILSLISGCVHSSCNPPDIASQAVDNVVLGKVVVAAAGNSGSSRTIGSPG 403
Qy 312 AASKVITGAVDSNDNINASFSGRPTA-----DGRLEKEPVVAPGVDIAPRASGYSM 363
Db 404 TARKATITGTSSTKSDIISWFSRGGVVMWMDAGITQALIKPVLVPG-----GTDS 454
Qy 364 GTP-----INDYTKASGTMATPHVSGVALILOAHPSWTPTALLETADIVAP 407
Db 455 GSEFCNPMEWFDNRICAMLNKEYLAISGTSMAITPLVSGAIALLKQKHPDWTPEIKGAV 514
Qy 408 IETADIVA--PKEIADIYAGRVNYKAIKDDVA 441
Db 515 KGTALNIGYDPE-----QAGRIINVRMIGLEERA 545

```

RESULT 5
US-11-433-614-14
Sequence 14, Application US/11433614
Publication No. US20060205019A1
GENERAL INFORMATION:
APPLICANT: Athena Biotechnologies, Inc.
APPLICANT: Menzel, Rolf
TITLE OF INVENTION: Methods and Compositions for Directed Gene Assembly
FILE REFERENCE: 46675-5004-01-US
CURRENT APPLICATION NUMBER: US/11/433,614
CURRENT FILING DATE: 2006-05-12
PRIOR APPLICATION NUMBER: US 09/920,118
PRIOR FILING DATE: 2001-07-31
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.3
SEQ ID NO 14
LENGTH: 379
TYPE: PR1
ORGANISM: Bacillus licheniformis
US-11-433-614-14

Query Match 16.2%; Score 557; DB 7; Length 379;
Best Local Similarity 34.3%; Pred. No. 1,8e-31;

Matches 150; Conservative 64; Mismatches 149; Indels 74; Gaps 14;

QY 3 GAKLLIIVLVGLVGSVAAAPKRYBYRNVKNTGLTTPGLFRKIQKLNPNRISTV 62
DB 10 GMTALMLVFTMAFSDSASAAQPAK-----NVEKDY-----I 41
QY 63 IVEFNH-REKEIAVRVLELMGAKRYVYVYHIIIPAIADLKVRDLVLSGLTGKAKLSGR 121
DB 42 VGFSSGVKTSVKKDIIKESGKVDKQFRIINAKALDKBALBEEV-----NDPVA 94
QY 122 FIOEDVYVTVSABIEGLDESAQAQVATYVNLGDSGSGITIGITDGDASHPDLOQVY 181
DB 95 YVEBDH--VAHALAQTPYGPILPKADKVOAQYKGNVAVAVLDGTIQSHHDL--NV 150
QY 182 GWVDFVNGRSPYRDDHGHGTHVASIAAGTGAASNGK--YKMAPGAKLAGIKVLGADSG 239
DB 151 GGAAPVAGAVNVDGNHGHV---AGTVAALDNTTGVAVAPNVLAVKVLNASSGSG 206
QY 240 SISTIIGVEMAVDNKDKYGIYINLSLIGSSQSSDGTSLISAQVNNAMADGIVCVAAQN 299
DB 207 SYSGIVSGIEMATTN---GMDVINSLG---GPSSTAMKQAVDVAHARGVVVAAGN 259
QY 300 SGP--NTYTGSPAAASKVITVGAVDSDNINASFSSRGPTADGRLEPVVAPGVDIAPR 357
DB 260 SGSSGNTNTIGYPAKYDSVIAVAVDPNSNRASFSSVG---AELEVAAPGAGVSTY 313
QY 358 ASGTSMGTPIINDYTTASGSMATPHVSGVGLILQAHPSWTDKTKTALIEADIVAPK 417
DB 314 PTST-----YATLNGTSMASPHVAGAAALILSKHPULASQVRNRLSTATYLG-- 362
QY 418 EIADIYAGAGRVNYYKA 434
DB 363 ---SSFYKGLINVEAA 377

RESULT 6
US-10-530-643-114
Sequence 114, Application US/10530643
Publication No. US2006025995A1
GENERAL INFORMATION:
APPLICANT: CAYOUETTE, Michelle
APPLICANT: HANSEN, Connie Jo
APPLICANT: MCCLORE, Amy
APPLICANT: SUN, May
APPLICANT: GRAMATIKOVA, Svetlana
APPLICANT: DYCAICO, Mark
APPLICANT: BARTON, Nelson R.
APPLICANT: STEGE, Justin T.
APPLICANT: ABOUSHADI, Nahla M.
TITLE OF INVENTION: PROTEASES, NUCLEIC ACIDS ENCODING THEM

TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
FILE REFERENCE: 564462004100
CURRENT APPLICATION NUMBER: US/10/530,643
CURRENT FILING DATE: 2006-04-07
PRIOR APPLICATION NUMBER: PCT/US03/32819
PRIOR FILING DATE: 2003-10-10
PRIOR APPLICATION NUMBER: 60/471,423
PRIOR FILING DATE: 2003-05-16
PRIOR APPLICATION NUMBER: 60/418,467
PRIOR FILING DATE: 2002-10-10
NUMBER OF SEQ ID NOS: 255
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 114
LENGTH: 1606
TYPE: PR1
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Obtained from an environmental sample.
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(32)
US-10-530-643-114

Query Match 15.7%; Score 539; DB 6; Length 1606;
Best Local Similarity 22.6%; Pred. No. 2.7e-29;

Matches 233; Conservative 111; Mismatches 235; Indels 454; Gaps 37;

QY 2 KGIKLLIIVLVGLV---GSVAAAPKRYBYRNVKNTGLTTPGLFRKIQKLNPNRISTV 45
DB 8 KFLSAFLTILVLSLVLPPNGTANALAKKOPKTPPODASALQQAALAQOTSLLGGR 67
QY 46 LFRKIQKLNPNRISTV-----VFNHREK 71
DB 68 LHKDILGSGSGEYVVIHLSEKAVGLBQGIHLKNGKMSQSEITLKKGKTQAQNTARK 127
QY 72 ELAVRVLBMGAKRY--VTH---IIPAIADLKVRDLVLSGLTGKAKLSGVRFQED 126
DB 128 EMTVK-----KVFEDVYSYDTVLNGLSGTVQANDLKL--LT-----ISGVYVSPD 173
QY 127 YKVTVSABIEGLDESAQAQVATYV---VNLGDSGSGITIGITDGDASHPDLOQ 178
DB 174 TTVYAOBGLQVDPKVDKMDTSLPFLGINKLMEBGIQGVAVAVLDGTIDASHPDKA 233
QY 179 KVIQWDFVNGRSPY---DDHG-----HGTVAASIAA 208
DB 224 YKGGKNTFVPHVDGYAPRADDSSTPLDRPANKREVENRGSSFTTGTHTVAGTIV 293
QY 209 GIGGAASNGKTYGMAPGAKLAGIKVLGADSGSISTIIKGVEMAVDNKDKYGIYINLSLG 268
DB 294 AOGNNEFG-IGIAFKVDLYSYRVLGAGSGATSGIITKADTAVIEK---IDVINLSLG 348
QY 269 SSQSSDGTSLISAQVNNAMADGIVCVAAAGSGPNTYVGSPPAAASKYITV----- 320
DB 349 GGANSE-IDGASPAIINNAMLAGTISVAVATGNSGRNKGMTGTPAISRLGIAVGNTTNBE 407
QY 321 ----- 320
DB 408 FOANVNVAGEFTYSKVNDLMAITFGKOVETOLAGYDLIAPGAGAKOYAADLVYTKV 467
QY 321 -----AVD----- 323
DB 468 ALVABSEIAPVDKVAALKKNGAVAIIVHNFAGTNAENKSDVFLGDAFEFIPTEFMSVTD 527
QY 324 -----SNDNTASFSSRGPTADGRLEPVVAPGVDIAPR 355
DB 528 GEALRAALAAPSTISPDQSSIKITGDDVNDSSRSPSTPNPDKDVAAPGNINISTI 587
QY 356 PRASGTSMGTPIINDYTTASGSMATPHVSGVGLILQAHPSWTDKTKTALIEADIVAPK 415
DB 588 PMYGNDEPPDADYSRAYSRTGTSMATPHIAGIALVAVQAMPVNAFVVKVLSATATVLD 647
QY 416 PEIADIYAGAGRVNYYKAIKYD----- 438

Db 648 TKKYDVFAQAGRVADYKARADVLAYALDTPASNDGTEVENIKGVTTFGPQKLDKNISVT 707
 QY 439 -----DVA-----KLTFTGVSAD----- 451
 Db 708 KKLNVDLKASAGDYVSVADYTKGPDGAKVTVDQSEFTLNGEQLLVNTLTASAELEKAD 767
 QY 452 -----KG-----SATHFDVSGATFVT-----ATLYV----- 473
 Db 768 ELIYGTHIKNGKDLSLPFAADGGAAYAVKDMETKTDLSPNGDVNDMLLYFTING 827
 QY 474 DPGSSIDLY-LYDPNGNEVDYSYTAHYGPEKGYNPN-----TAGWTW-----KVSXKG 524
 Db 828 DVGFTNIEIEMDMNPTGKXSDY-----IGYHASPDLTLAGSYOLPLLEKTYQWOG 879
 QY 525 AANYQDVVSDGSL-----SOSGGGNPNPNP-NPPTPTDQ-TFTGSVND----- 569
 Db 880 TMS--EVLPGDLTYTIDFSAEYSGNPKTIADYGVVVVXKSTAGTLEGAVADGKTGKIT 936
 QY 570 --YMD-----TSPTFTMNVN-----SGATKITGD-----LTPTSY 598
 Db 937 DKYVDQKELVYKGMGYDLNLTUKSATYEVYANDAVVDSGAVKLAQDGTAFDLPTFDKX 996
 QY 599 NDLDLYLDPNGN 611
 Db 997 NNVTYKADAGN 1009

RESULT 7
 US-10-532-605-2
 ; Sequence 2, Application US/10532605
 ; Publication No. US20060134092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Miwa, Takehiro
 ; TITLE OF INVENTION: Method for digesting proteins highly resistant to denaturation
 ; FILE REFERENCE: 087625
 ; CURRENT APPLICATION NUMBER: US/10/532,605
 ; CURRENT FILING DATE: 2005-04-25
 ; PRIOR APPLICATION NUMBER: PCT/JP2003/013658
 ; PRIOR FILING DATE: 2003-10-24
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 2
 ; LENGTH: 274
 ; TYPE: PRT
 ; ORGANISM: Bacillus licheniformis
 ; US-10-532-605-2

Query Match 15.5%; Score 532; DB 6; Length 274;
 Best Local Similarity 44.6%; Pred. No. 6.6e-30; Indels 36; Gaps 9;
 Matches 131; Conservative 37; Mismatches 90;

QY 145 VMAATVYVNLGDSGITTIGIIDTGDASHPDLOGKVIQWVDFVNGRSGYEDHGHGTHVA 204
 Db 11 IKADKYQAGQYKGANVKGIIIDTGIASHTDL--KVGASAFVSGESYNTDNGHGHGTHV- 67
 QY 205 STAACTGAASNGK--YKMAPGAKLAGIYVLGADSGSISTIIKGYEAWVDNKKYKIV 262
 Db 68 ---AGTVALDNTTGTGLVAPNVLAIKVLNSSGSGTSAISGEMATQN---GLDV 120
 QY 263 INLSTGSSSSDQDBLSQAVNNAMAGIYVCAANSNP--NTYVGSPPAASKIYTYG 320
 Db 121 INNSTL--GPSGSTALKQAVDAVASGIWVAANAASSGSGQNTIGYPAKYDSYIANG 177
 QY 321 AVDSNDINIASFSRGTADGRLEPEVAVAGVDIIAPRAGTSMGTPIINDYTYKASGTSMA 380
 Db 178 AVDSNNGRASFSVGS-----ELVMAAGVSVSYSTPNSMT-----TYSINQTSMA 222
 QY 381 TPRVSGVALLIQAHSWTPDKYKTAIIETADIAPKEIADIAYGGRVNVYKA 434
 Db 223 SPHVAGAAALLISKYPTLSASQVRNRLSSTATNLGD-----SFYGGGLINVEAA 272

RESULT 8
 US-11-452-695-2
 ; Sequence 2, Application US/11452695
 ; Publication No. US20060222641A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Neil H. Riordan
 ; TITLE OF INVENTION: METHOD AND COMPOSITION FOR PREVENTING OR
 ; FILE REFERENCE: AIDAN.005DV1
 ; CURRENT APPLICATION NUMBER: US/11/452,695
 ; CURRENT FILING DATE: 2006-06-14
 ; PRIOR APPLICATION NUMBER: 60/468948
 ; PRIOR FILING DATE: 2003-05-07
 ; PRIOR APPLICATION NUMBER: 10/647131
 ; PRIOR FILING DATE: 2003-08-22
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 381
 ; TYPE: PRT
 ; ORGANISM: Bacillus Subtilis
 ; US-11-452-695-2

Query Match 15.4%; Score 527; DB 7; Length 381;
 Best Local Similarity 37.9%; Pred. No. 2.4e-29;
 Matches 145; Conservative 54; Mismatches 126; Indels 58; Gaps 15;

QY 61 TVIVFBNREKEIAVRVLELMGAKRYRYTHIIPALAADL--KRDVLVISGLTGKAKL 117
 Db 46 TWSAMSARKKD---VISEKGGKQKQKYNVAAAATLDEKAVKEL-----KRD 91
 QY 118 SGVRFIOEDYKVTSAE-LEGDESAQVMAVYVNNLGYDGSIGTIGIIDTGDASHPDL 176
 Db 92 PAVAYVEEDH---IAHEVAQSVPYGISQIKAPALHSQGTGSNVKAVAVDSIGDSHPDL 148
 QY 177 QCKVIGWDFVNGRSPYD--HGHTHVASIAGTGAASNGK--YKMAPGAKLAGIKYL 233
 Db 149 --NVAGASFPVSENPYDQSSHGTHV---AGTIALNNSIGVLGVAPASLAVVKVL 202
 QY 234 GADSGSISTIIKGYEAWVDNKKYKIKYINLSLSSGSSDGTSLSQAVNNAMAGIYV 293
 Db 203 DSTGSGQYSWIINGEMALSN---NMDVYNNSLG---GPTSTALKTVVDRAVSSGIYV 255
 QY 294 CVAAGNSGP--NTYVGSPPAASKIYITVGAVDNSNDINIASFSRGTADGRLEPEVAPGV 351
 Db 256 AAAAGNBSGSGSTVGYPAKYPTIAGAVANSSNORASFSSVSELD-----VMAPEV 309
 QY 352 DIAPRASGTSMTPIINDYTYKASGTSMAITPHVSGVALLIQAHSWTPDKYKTALETA 411
 Db 310 SIQSTLPQGT-----YGAVNGTSMATPHVAGAAALLISKHPWTYTAQVRDRLESTA 360
 QY 412 DIVAPKEIADIAYGGRVNVYKA 434
 Db 361 TVLG-----SSFYGGGLINVOAA 379

RESULT 9
 US-10-541-737-3
 ; Sequence 3, Application US/10541737
 ; Publication No. US20060252155A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Leeflang, Chris
 ; APPLICANT: Van der Kleij, Wilhelmus A.H.
 ; TITLE OF INVENTION: Methods for Site-Directed Mutagenesis
 ; FILE REFERENCE: GC787-2
 ; CURRENT APPLICATION NUMBER: US/10/541,737
 ; CURRENT FILING DATE: 2005-07-08
 ; PRIOR APPLICATION NUMBER: PCT/US04/01334
 ; PRIOR FILING DATE: 2004-01-16
 ; PRIOR APPLICATION NUMBER: US 60/440,792
 ; PRIOR FILING DATE: 2003-01-16
 ; NUMBER OF SEQ ID NOS: 6

SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 274
 TYPE: PR1
 ORGANISM: Bacillus licheniformis
 US-10-541-737-3

Query Match 15.3%, Score 525, DB 6, Length 274;
 Best Local Similarity 43.5%, Pred. No. 2, le-29;
 Matches 126; Conservative 40; Mismatches 90; Indels 36; Gaps 10;

QY 145 VMATYVWNLGYDGGITITIGTIDGASHPDLOGKVIQWDFVNGRSPYDDHGHGTHVA 204
 DB 11 IKADKVAQGGKGVANVAVLDTGISOASHPDL--NVVGAGSFVAGFAVNTDGNHGHGTHV- 67
 QY 205 SIAGTGAAGSNGK--YKGMAPGAKLAGIKVLGADGSSISITIIKGVMAVDNKKYKIXV 262
 DB 68 ---AGTVAALDNTTGVLVAVPSVSLVAVKVLNLSGSSGYSIGVIGEMATTN---GMDV 120
 QY 263 INTSLGSSQSSDGTDSLQAVNNMADGIWCVVAGNSG--PNTYVGSPPAASKVITVG 320
 DB 121 INMSLG---GASGSTAKQAVDNAVARGVVAAGNSGSGTNTIGYPAKXDSVIAVG 177
 QY 321 AVDSNDVIASFSSRGPTADGRLKPEVVAPEVDIIAPRAGTSGTPIINDYTTAAGTSM 380
 DB 178 AVDSNSNRASFSSVG-----ABLEVVAPE-----AGVYSTYPTNTYAT-LNGTSM 222
 QY 381 TPRHVGVALILOAHPSMTPEKVKYKALITETADIAPKEIDIAVAGARVNVYKA 434
 DB 223 SPHVAGAAAILLSGHPULSASQVRNRLSTATYIG---SFFYKGLINVEAA 272

RESULT 10
 US-11-452-695-3

Sequence 3, Application US/11452695
 Publication No. US20060222641A1
 GENERAL INFORMATION:
 APPLICANT: Neill H. Riordan
 TITLE OF INVENTION: METHOD AND COMPOSITION FOR PREVENTING OR
 FILE REFERENCE: AIDAN.005DVL
 CURRENT APPLICATION NUMBER: US/11/452,695
 PRIOR FILING DATE: 2006-06-14
 PRIOR APPLICATION NUMBER: 60/468948
 PRIOR FILING DATE: 2003-05-07
 PRIOR APPLICATION NUMBER: 10/647131
 NUMBER OF SEQ ID NOS: 7
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 381
 TYPE: PR1
 ORGANISM: Bacillus Subtilis
 US-11-452-695-3

Query Match 15.3%, Score 524, DB 7, Length 381;
 Best Local Similarity 37.9%, Pred. No. 3, le-29;
 Matches 145; Conservative 53; Mismatches 127; Indels 58; Gaps 15;

QY 61 TVIVFENHREKEIAVAVLELMGAKVRYVYHIIIPAIADL---KYRDLVLSGLTGKAKL 117
 DB 46 TWSAASSAKKD---VISEKGVQKQKRYVNAATAITLDEKAVKEL-----KKD 91
 QY 118 SGVRFIOEDYKVTVAE-LEGIDESAAQVMAVYVNNLGYDGGITITIGTIDGASHPDL 176
 DB 92 PSVAAYVEEDH---IAHEVAQSVPGISQIKAPALHSOGTGSNVKVAVIDSGIDSHPDL 148
 QY 177 QGKVIQWDFVNGRSPYDD-HGHGTHVASIAAGTGAASNGK--YKGMAPGAKLAGIKVL 233
 DB 149 --NVRGASFPVSEINPYDODSSHGTHV---AGTIALNNSIGVLGVAAPSASLYAVKVL 202
 QY 234 GADGSGSISTIIKGVEMAVDNKKYKIGIKVNLISGSSQSSDGTDSLQAVNNMADGIW 293

DB 203 DSTGSGQSWIINGIEMALSN---NMDVINMSLG---GPTGSTALKTVDKAVSSGIWV 255
 QY 294 CVAAGNSGP--NTYVGSPPAASKVITVGAVDSDNDVIASFSSRGPTADGRLKPEVVAPEV 351
 DB 256 AAAAGNEGSGSISTVGYPAKYPSTIAVGAVNSNORAFSSVSGSLD-----VNAPEV 309
 QY 352 DIAPRAGTSMGTPINDYTTKASGTSMATPHVSGVALILOAHPSMTPEKVKYKALITETA 411
 DB 310 SIQSTLPQGT-----YGAVNGTSMATPHVAGAAAILLSKHPYVTAQVRDLRESTA 360
 QY 412 DIVAPKEIADIAGARVNVYKA 434
 DB 361 TYLG---NSFFYKGLINVOAA 379

RESULT 11

US-11-452-695-4
 Sequence 4, Application US/11452695
 Publication No. US20060222641A1
 GENERAL INFORMATION:
 APPLICANT: Neill H. Riordan
 TITLE OF INVENTION: METHOD AND COMPOSITION FOR PREVENTING OR
 FILE REFERENCE: AIDAN.005DVL
 CURRENT APPLICATION NUMBER: US/11/452,695
 PRIOR FILING DATE: 2006-06-14
 PRIOR APPLICATION NUMBER: 60/468948
 PRIOR FILING DATE: 2003-05-07
 PRIOR APPLICATION NUMBER: 10/647131
 NUMBER OF SEQ ID NOS: 7
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 4
 LENGTH: 381
 TYPE: PR1
 ORGANISM: Bacillus Subtilis
 US-11-452-695-4

Query Match 15.3%, Score 524, DB 7, Length 381;
 Best Local Similarity 37.9%, Pred. No. 3, le-29;
 Matches 145; Conservative 53; Mismatches 127; Indels 58; Gaps 15;

QY 61 TVIVFENHREKEIAVAVLELMGAKVRYVYHIIIPAIADL---KYRDLVLSGLTGKAKL 117
 DB 46 TWSAASSAKKD---VISEKGVQKQKRYVNAATAITLDEKAVKEL-----KKD 91
 QY 118 SGVRFIOEDYKVTVAE-LEGIDESAAQVMAVYVNNLGYDGGITITIGTIDGASHPDL 176
 DB 92 PSVAAYVEEDH---IAHEVAQSVPGISQIKAPALHSOGTGSNVKVAVIDSGIDSHPDL 148
 QY 177 QGKVIQWDFVNGRSPYDD-HGHGTHVASIAAGTGAASNGK--YKGMAPGAKLAGIKVL 233
 DB 149 --NVRGASFPVSEINPYDODSSHGTHV---AGTIALNNSIGVLGVAAPSASLYAVKVL 202
 QY 234 GADGSGSISTIIKGVEMAVDNKKYKIGIKVNLISGSSQSSDGTDSLQAVNNMADGIW 293
 DB 203 DSTGSGQSWIINGIEMALSN---NMDVINMSLG---GPTGSTALKTVDKAVSSGIWV 255
 QY 294 CVAAGNSGP--NTYVGSPPAASKVITVGAVDSDNDVIASFSSRGPTADGRLKPEVVAPEV 351
 DB 256 AAAAGNEGSGSISTVGYPAKYPSTIAVGAVNSNORAFSSVSGSLD-----VMAPEV 309
 QY 352 DIAPRAGTSMGTPINDYTTKASGTSMATPHVSGVALILOAHPSMTPEKVKYKALITETA 411
 DB 310 SIQSTLPQGT-----YGAVNGTSMATPHVAGAAAILLSKHPYVTAQVRDLRESTA 360
 QY 412 DIVAPKEIADIAGARVNVYKA 434
 DB 361 TYLG---NSFFYKGLINVOAA 379

RESULT 12

US-11-433-614-16

Sequence 16, Application US/11433614
Publication No. US20060205019A1
GENERAL INFORMATION:
APPLICANT: Athena Biotechnologies, Inc.
TITLE OF INVENTION: Method and Compositions for Directed Gene Assembly
FILE REFERENCE: 46675-5004-01-US
CURRENT APPLICATION NUMBER: US/11/433,614
CURRENT FILING DATE: 2006-05-12
PRIOR APPLICATION NUMBER: US 09/920,118
PRIOR FILING DATE: 2001-07-31
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.3
SEQ ID NO 16
LENGTH: 381
TYPE: PRF
ORGANISM: Bacillus subtilis
US-11-433-614-16

Query Match 15.1%; Score 518; DB 7; Length 381;
Best Local Similarity 37.6%; Pred. No. 1e-28;
Matches 144; Conservative 53; Mismatches 128; Indels 58; Gaps 15;

QY 61 TVIVENHREKEIAVLELMGAKVRYVYHIIIPALADI--KVRDLVIGLGGKAYL 117
DB 46 TWSAASAKKCD---VISEKGKVKQKPKYVAAATLDEKAVKEL-----KKD 91
QY 118 SGVRPIQEDYKXTVSAE-LEGIDESAAQVMATYVNVLYGDSGITITIGTIDGASHPTL 176
DB 92 PSVAAYVEEH---IAHEVYQSVYVGIQIKAPALHSQYTSNVKAVAVIDSGIDSSHPTL 148
QY 177 QGKVIQWVFNVRGRSPYD-HGHGTHVASIAAGTGAANGK--YKGMAPGAKLAGIKYL 233
DB 149 --NMGASFSVSETPYDGGSHGTHV---AGTIALNNSIGVLGVAAPSILAVKYL 202
QY 234 GADSGSISTIIKGVEMAVDNKDKXIKYINLSLSSGSSDGTSLISQAVNNAMPAGIYV 293
DB 203 DSTGSGQYWIINGIEMALSN---NMDVYINMSLG---GPTGSLTKTVVDKAVSGIYV 255
QY 294 CVAANGSG--NTYTGSPAAASKVTYTGAVDSNDNINAFSSRGPRADRLKPEVAVPY 351
DB 256 AAAAGSGSGSTVGYPAKYPSTIYAVGAVNNSNORASFSSAGSELD-----VMAPEV 309
QY 352 DIIAPRAGTSMGTPIINDYTKASGSMATPHVSGGALLIQAPSWTPDKYKTLIETA 411
DB 310 SIQSTLPGST-----YGAINGSMATPHVAAAALILSKHPYTNAAQVDRLESTA 360
QY 412 DIVAPKEIADIYAGARVNVYKA 434
DB 361 TYLG---NSFYGKGLINVQAA 379

RESULT 13
US-10-567-073-3
Sequence 3, Application US/10567073
Publication No. US20060134740A1
GENERAL INFORMATION:
APPLICANT: Bryan, Philip N.
TITLE OF INVENTION: Engineered Proteases for Affinity Purification and Processing of
FILE REFERENCE: 4115-181
CURRENT APPLICATION NUMBER: US/10/567,073
CURRENT FILING DATE: 2006-02-03
PRIOR APPLICATION NUMBER: US 60/493,032
PRIOR FILING DATE: 2003-08-06
PRIOR APPLICATION NUMBER: PCT/US04/021049
PRIOR FILING DATE: 2004-06-29
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.3
SEQ ID NO 3
LENGTH: 275
TYPE: PRF
ORGANISM: Bacillus amyloliquefaciens

US-10-567-073-3

Query Match 15.1%; Score 517.5; DB 6; Length 275;
Best Local Similarity 43.3%; Pred. No. 7e-29;
Matches 129; Conservative 41; Mismatches 89; Indels 39; Gaps 11;

QY 143 AQWATYVWNLGYDGGSGITITIGTIDGASHPTLQKVIQWVFNVRGRSPY-DHGHGT 201
DB 9 SQIKAPALHSQYTSNVKAVAVIDSGIDSSHPTL--KVAAGASWVPSFTNPQDNNSHOT 66
QY 202 HVASIAAGTGAASNGK--YKGMAPGAKLAGIYVLCADSGSISTIIKGVEMAVDNKDKKG 259
DB 67 HV---AGTVAALNNSIGVLGVAAPSILYAVKVLGADSGQYWIINGIEMALSN---N 118
QY 260 IKVINLSLSSGSSDGTSLISQAVNNAMPAGIYVCAANGSGP--NTYTGSPAAASKYI 317
DB 119 MDVINMSLG---GPGSAAALKAAYDKAASGVVVAALAGNBSSTSTVGTGKPPSYI 175
QY 318 TVGAVDSNDNINAFSSRGPRADRLKPEVAVPGVDIIAPRAGTSMGTPIINDYTKASGT 377
DB 176 AVGAVDSSNQRASFSSVGPSELD-----VMAPEVSI-----QSTLPGRKYGAVNGT 220
QY 378 SMATPHVSGVGAALLIQAPSWTPDKYKTLIETADIAPKEIAD-IAYGAGRNVYKA 434
DB 221 SWSAPHVAGAAALILSKHPNWTNTQVRSSLENTT-----TKLGDSFYGKGLINVQAA 273

RESULT 14

US-10-541-737-1
Sequence 1, Application US/10541737
Publication No. US20060252155A1
GENERAL INFORMATION:
APPLICANT: LeeJiang, Chris
TITLE OF INVENTION: Methods for Site-Directed Mutagenesis
FILE REFERENCE: GC787-2
CURRENT APPLICATION NUMBER: US/10/541,737
CURRENT FILING DATE: 2005-07-08
PRIOR APPLICATION NUMBER: PCT/US04/01334
PRIOR FILING DATE: 2004-01-16
PRIOR APPLICATION NUMBER: US 60/440,792
PRIOR FILING DATE: 2003-01-16
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 275
TYPE: PRF
ORGANISM: Bacillus amyloliquefaciens
US-10-541-737-1

Query Match 15.1%; Score 517.5; DB 6; Length 275;
Best Local Similarity 43.3%; Pred. No. 7e-29;
Matches 129; Conservative 41; Mismatches 89; Indels 39; Gaps 11;

QY 143 AQWATYVWNLGYDGGSGITITIGTIDGASHPTLQKVIQWVFNVRGRSPY-DHGHGT 201
DB 9 SQIKAPALHSQYTSNVKAVAVIDSGIDSSHPTL--KVAAGASWVPSFTNPQDNNSHOT 66
QY 202 HVASIAAGTGAASNGK--YKGMAPGAKLAGIYVLCADSGSISTIIKGVEMAVDNKDKKG 259
DB 67 HV---AGTVAALNNSIGVLGVAAPSILYAVKVLGADSGQYWIINGIEMALSN---N 118
QY 260 IKVINLSLSSGSSDGTSLISQAVNNAMPAGIYVCAANGSGP--NTYTGSPAAASKYI 317
DB 119 MDVINMSLG---GPGSAAALKAAYDKAASGVVVAALAGNBSSTSTVGTGKPPSYI 175
QY 318 TVGAVDSNDNINAFSSRGPRADRLKPEVAVPGVDIIAPRAGTSMGTPIINDYTKASGT 377
DB 176 AVGAVDSSNQRASFSSVGPSELD-----VMAPEVSI-----QSTLPGRKYGAVNGT 220
QY 378 SMATPHVSGVGAALLIQAPSWTPDKYKTLIETADIAPKEIAD-IAYGAGRNVYKA 434

DB 221 SMASPHVAGAAALILSKHPMTNTQVRSILENTT-----TKLGDSFYKGLINVQAA 273

RESULT 15

US-10-526-324-20

Sequence 20, Application US/10526324

Publication No. US20060248617A1

GENERAL INFORMATION:

APPLICANT: Imanaka, Takayuki

APPLICANT: Atomi, Haruyuki

TITLE OF INVENTION: METHOD OF TARGETED GENE DISRUPTION, GENOME OF

TITLE OF INVENTION: HYPERHERMOSTABLE BACTERIUM AND GENOME CHIP USING

TITLE OF INVENTION: THE SAME

FILE REFERENCE: 490051.40USPC

CURRENT APPLICATION NUMBER: US/10/526,324

PRIOR FILING DATE: 2005-02-28

PRIOR APPLICATION NUMBER: PCT/IB2003/003597

PRIOR FILING DATE: 2003-08-29

PRIOR APPLICATION NUMBER: JP 2002-319011

PRIOR FILING DATE: 2002-08-30

NUMBER OF SEQ ID NOS: 2167

SOFTWARE: PatentIn version 3.1

SEQ ID NO 20

LENGTH: 524

TYPE: PRT

ORGANISM: Thermococcus kodakarensis KOD1

FEATURE:

NAME/KEY: misc feature

LOCATION: (414542)..(414542)

OTHER INFORMATION: n is a or c or g or t.

FEATURE:

NAME/KEY: misc feature

LOCATION: (786890)..(786890)

OTHER INFORMATION: n is a or c or g or t.

FEATURE:

NAME/KEY: misc feature

LOCATION: (786907)..(786907)

OTHER INFORMATION: n is a or c or g or t.

FEATURE:

NAME/KEY: misc feature

LOCATION: (786944)..(786944)

OTHER INFORMATION: n is a or c or g or t.

FEATURE:

NAME/KEY: misc feature

LOCATION: (839139)..(839139)

OTHER INFORMATION: n is a or c or g or t.

FEATURE:

NAME/KEY: misc feature

LOCATION: (1128488)..(1128488)

OTHER INFORMATION: n is a or c or g or t.

FEATURE:

NAME/KEY: misc feature

LOCATION: (1128499)..(1128499)

OTHER INFORMATION: n is a or c or g or t.

FEATURE:

NAME/KEY: misc feature

LOCATION: (1128505)..(1128505)

OTHER INFORMATION: n is a or c or g or t.

FEATURE:

NAME/KEY: misc feature

LOCATION: (1128517)..(1128517)

OTHER INFORMATION: n is a or c or g or t.

FEATURE:

NAME/KEY: misc feature

LOCATION: (1128539)..(1128539)

OTHER INFORMATION: n is a or c or g or t.

FEATURE:

NAME/KEY: misc feature

LOCATION: (1561400)..(1561400)

OTHER INFORMATION: n is a or c or g or t.

FEATURE:

NAME/KEY: misc feature

LOCATION: (1561477)..(1561477)

OTHER INFORMATION: n is a or c or g or t.

FEATURE:

NAME/KEY: misc feature

LOCATION: (1561545)..(1561545)

OTHER INFORMATION: n is a or c or g or t.

FEATURE:

NAME/KEY: misc feature

LOCATION: (1767941)..(1767941)

OTHER INFORMATION: n is a or c or g or t.

US-10-526-324-20

Query Match 14.7%; Score 504.5; DB 6; Length 524;

Best Local Similarity 35.9%; Pred. No. 1.5e-27;

Matches 157; Conservative 67; Mismatches 154; Indels 59; Gaps 18;

QY 51 OKLNPEBISTIVIF---ENHREKEIAVAVLEIMGAKVYVYHIIIPAIADLKVRDLVI 107

DB 104 QKVONDEPIQWIFITTKPRNDLKE---ALEHLAGVLYVDDDIYALIS--LKARPSL-I 156

QY 108 SGLTGKAKLDSVRF-IQEDY-----KTVSAR-LRGLDSAAQ 144

DB 157 KNLVVSQAFDPDYRFIRVWPDLYETGPTILENETFTPANVITTPQVSGAIVNIXL 216

QY 145 VVATTYVNL-GYDGSITIGIIDTGIDASHPDLOGKVIGMVDVFNKRSIPYDD-HGHTG 202

DB 217 VKADILMSKRGITGKGVVAVLDTGVDCHVMLOGACVGFENVTDE--PAKDNLNGHT 274

QY 203 VASIAAG--TGAASNGK---YKGMAPGAKLAGIKVLGADSGSISTIIKGVEMAVDNKDK 257

DB 275 VAGIINGRPPTKYTWBKEVYVSGVAPEANILAVKVLGQGGGTMTIIGLDIVVWKKK 334

QY 258 YGIR--VINLSLSSQSOSDGTDSLQAANNMADGIVCVAAAGNSGPNITYVGSPPAAAK 315

DB 335 HPEEPVIVMSLGSPPGSPRDPVQVBOIIREHLPVILAGN---BFAVIDSPGIATG 391

QY 316 VITVGAVDSDNINIASFSRGPTRD-GRLEPVVAPGVDTIAPRAGTSMGTPIINDYTTA 374

DB 392 AITVAAVDNMMKVASFSGKPGIINIVDKPIDIAAPVKILSAXA---GT--RNEFIAM 444

QY 375 SGTSMATPHVSGVAGALILAHPSWTPDKVTKLILETADIVAPRSIADIVAGRVNVYKA 434

DB 445 SGTSMATPHVSGVVALILQHGDLTPETIKLILEKTAIYPLDGDILALPTMSGAGVDAIYAA 504

QY 435 IKYDYAKLTFGTSVAD 451

DB 505 VK---ABPSBSGGLTD 517

Search completed: January 6, 2007, 22:38:59

Job time : 33 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2007 Bioacceleration Ltd.

OW protein - protein search, using sw model

Run on: January 6, 2007, 22:31:42 ; Search time 36 Seconds

(Without alignments)
1602.297 Million cell updates/sec

Title: US-10-800-684-1

Perfect score: 3437

Sequence: 1 MKRLGAVLVAVLVLACT.....YAVSTYGMADYOLKAVVYVG 659

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:
1: /EMC_Celerra_SIDS3/prodata/2/1aa/5 COMB.pep:*
2: /EMC_Celerra_SIDS3/prodata/2/1aa/6 COMB.pep:*
3: /EMC_Celerra_SIDS3/prodata/2/1aa/7 COMB.pep:*
4: /EMC_Celerra_SIDS3/prodata/2/1aa/H_COMB.pep:*
5: /EMC_Celerra_SIDS3/prodata/2/1aa/PCTUS_COMB.pep:*
6: /EMC_Celerra_SIDS3/prodata/2/1aa/RE_COMB.pep:*
7: /EMC_Celerra_SIDS3/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------------|
| 1 | 3437 | 100.0 | 659 | 2 | US-08-894-818B-1 |
| 2 | 3437 | 100.0 | 659 | 2 | US-09-445-472-12 |
| 3 | 3437 | 100.0 | 659 | 2 | US-10-090-624-12 |
| 4 | 3437 | 100.0 | 659 | 2 | US-09-841-553-1 |
| 5 | 2914 | 84.8 | 659 | 2 | US-08-894-818B-5 |
| 6 | 2914 | 84.8 | 659 | 2 | US-09-841-553-5 |
| 7 | 2483.5 | 72.3 | 654 | 2 | US-08-894-818B-35 |
| 8 | 2483.5 | 72.3 | 654 | 2 | US-09-445-472-16 |
| 9 | 2483.5 | 72.3 | 654 | 2 | US-10-090-624-16 |
| 10 | 2483.5 | 72.3 | 654 | 2 | US-09-841-553-35 |
| 11 | 2138.5 | 62.2 | 522 | 2 | US-08-894-818B-3 |
| 12 | 2138.5 | 62.2 | 522 | 2 | US-09-445-472-4 |
| 13 | 2138.5 | 62.2 | 522 | 2 | US-10-090-624-4 |
| 14 | 2138.5 | 62.2 | 522 | 2 | US-09-841-553-3 |
| 15 | 1707 | 49.7 | 412 | 2 | US-09-445-472-1 |
| 16 | 1707 | 49.7 | 412 | 2 | US-10-090-624-1 |
| 17 | 1148 | 33.4 | 237 | 1 | US-08-750-532-18 |
| 18 | 665 | 19.3 | 734 | 2 | US-09-000-016-4 |
| 19 | 665 | 19.3 | 734 | 2 | US-09-514-340-4 |
| 20 | 665 | 19.3 | 823 | 2 | US-09-000-016-2 |
| 21 | 665 | 19.3 | 823 | 2 | US-09-514-340-2 |
| 22 | 647.5 | 18.8 | 520 | 2 | US-09-000-016-7 |
| 23 | 647.5 | 18.8 | 520 | 2 | US-09-514-340-7 |
| 24 | 560.5 | 16.3 | 379 | 1 | US-08-685-774-2 |
| 25 | 557.5 | 16.2 | 350 | 1 | US-07-923-260A-4 |
| 26 | 557.5 | 16.2 | 379 | 1 | US-08-845-295A-1 |

| | | | | | | |
|----|-------|------|------|---|------------------|--------------------|
| 27 | 557.5 | 16.2 | 379 | 2 | US-09-140-933-1 | Sequence 1, Appli |
| 28 | 557.5 | 16.2 | 379 | 2 | US-09-146-661-1 | Sequence 1, Appli |
| 29 | 557.5 | 16.2 | 379 | 2 | US-09-150-515-1 | Sequence 1, Appli |
| 30 | 553 | 16.1 | 378 | 1 | US-07-772-087-4 | Sequence 4, Appli |
| 31 | 535.5 | 15.6 | 1398 | 1 | US-08-750-532-9 | Sequence 9, Appli |
| 32 | 535.5 | 15.6 | 1398 | 2 | US-08-894-818B-8 | Sequence 8, Appli |
| 33 | 535.5 | 15.6 | 1398 | 2 | US-09-445-472-6 | Sequence 6, Appli |
| 34 | 535.5 | 15.6 | 1398 | 2 | US-10-090-624-6 | Sequence 6, Appli |
| 35 | 535.5 | 15.5 | 1398 | 2 | US-09-841-553-8 | Sequence 8, Appli |
| 36 | 533.5 | 15.5 | 352 | 1 | US-07-923-260A-2 | Sequence 2, Appli |
| 37 | 533.5 | 15.5 | 382 | 2 | US-09-255-502-2 | Sequence 2, Appli |
| 38 | 533.5 | 15.5 | 382 | 2 | US-09-445-472-31 | Sequence 31, Appli |
| 39 | 533.5 | 15.5 | 382 | 2 | US-09-659-749-2 | Sequence 2, Appli |
| 40 | 533.5 | 15.5 | 382 | 2 | US-09-672-105-2 | Sequence 2, Appli |
| 41 | 533.5 | 15.5 | 382 | 2 | US-10-090-624-31 | Sequence 31, Appli |
| 42 | 533.5 | 15.5 | 382 | 2 | US-10-104-693-2 | Sequence 2, Appli |
| 43 | 532.5 | 15.5 | 382 | 2 | US-09-178-155-2 | Sequence 2, Appli |
| 44 | 532.5 | 15.5 | 382 | 2 | US-09-178-173A-2 | Sequence 2, Appli |
| 45 | 532.5 | 15.5 | 382 | 2 | US-09-177-353-2 | Sequence 2, Appli |

ALIGNMENTS

RESULT 1
US-08-894-818B-1
Sequence 1, Application US/08894818B
Patent No. 6261822
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MITTA, Masanori
APPLICANT: ASADA, Kiyoko
APPLICANT: TSUNASAWA, Susumu
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894, 818B
FILING DATE: 20-MAY-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03253
FILING DATE: 07-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25, 618
REFERENCE/DOCKET NUMBER: TAKAKURA=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

APP-THIS COPY
ANV

MOLECULE TYPE: peptide
US-08-894-8188-1

Query Match 100.0%; Score 3437; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 1.9e-236;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MKRLGAVVLAIVLVLGAGTALAAPVKPVRRNNAVQOKYGLTLPGLFKVQRMNNQEV 60
DB 1 MKRLGAVVLAIVLVLGAGTALAAPVKPVRRNNAVQOKYGLTLPGLFKVQRMNNQEV 60
QY 61 DTVIMFGSYGDRRAVAVKRLMGAQVYKYSKIIIPAAVKIKARDLLIAGMIDTGYFGNT 120
DB 61 DTVIMFGSYGDRRAVAVKRLMGAQVYKYSKIIIPAAVKIKARDLLIAGMIDTGYFGNT 120
QY 121 RVSGIKFIQEDYKQVVDATSVSQIGADTVMNSLGVDGSGVVAIVDTGIDANHPDLK 180
DB 121 RVSGIKFIQEDYKQVVDATSVSQIGADTVMNSLGVDGSGVVAIVDTGIDANHPDLK 180
QY 181 VIGMYDAVNGRSTPYDDQGHGTHVAGIVAGTGSVNSQYIGVAPGAKLVGVKVLGADGSGS 240
DB 181 VIGMYDAVNGRSTPYDDQGHGTHVAGIVAGTGSVNSQYIGVAPGAKLVGVKVLGADGSGS 240
QY 241 VSTIIAGVDVWVQNKDKYGRVNLISGSSQSSDGTDSLQAVNNAMDAGIIVCVAAGNS 300
DB 241 VSTIIAGVDVWVQNKDKYGRVNLISGSSQSSDGTDSLQAVNNAMDAGIIVCVAAGNS 300
QY 301 GPNTYVGSPPAAASKVITVGAVDSDNINAFSSRGPTADRLKPEVVA PGVDIIAPRAGS 360
DB 301 GPNTYVGSPPAAASKVITVGAVDSDNINAFSSRGPTADRLKPEVVA PGVDIIAPRAGS 360
QY 361 TSMGTPINDYTTKASGTSMATPHVSGVGLIIQAHPSMTDPKVTALIFRADIAPRKEIA 420
DB 361 TSMGTPINDYTTKASGTSMATPHVSGVGLIIQAHPSMTDPKVTALIFRADIAPRKEIA 420
QY 421 DIAYGARNVNVYKAIKYDDYAKLFTFGSVADKGSATHTFVSQATFTATILYMDTGSDDI 480
DB 421 DIAYGARNVNVYKAIKYDDYAKLFTFGSVADKGSATHTFVSQATFTATILYMDTGSDDI 480
QY 481 DLVLYDPNGNEVDYSTAYYGFKEVGYNPTAGTWTYKVVSYKGAANYQVDVSDGSLSQ 540
DB 481 DLVLYDPNGNEVDYSTAYYGFKEVGYNPTAGTWTYKVVSYKGAANYQVDVSDGSLSQ 540
QY 541 SGGNNPNPNPNPPTPTDTQTFTGSVNDYMDTSDFTMNVNSGATKLTGDLTFDTSYND 600
DB 541 SGGNNPNPNPNPPTPTDTQTFTGSVNDYMDTSDFTMNVNSGATKLTGDLTFDTSYND 600
QY 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFLLVYASTYGMADYQLKAVVYVG 659
DB 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFLLVYASTYGMADYQLKAVVYVG 659
```

RESULT 2

US-09-445-472-12
Sequence 12, Application US/09445472
Patent No. 6358726
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOTO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshi
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/09/445,472
CURRENT FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patentin version 3.0
SEQ ID NO 12
LENGTH: 659
TYPE: PRT

No BLAST part
problem

ORGANISM: Thermococcus celer
US-09-445-472-12

Query Match 100.0%; Score 3437; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 1.9e-236;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MKRLGAVVLAIVLVLGAGTALAAPVKPVRRNNAVQOKYGLTLPGLFKVQRMNNQEV 60
DB 1 MKRLGAVVLAIVLVLGAGTALAAPVKPVRRNNAVQOKYGLTLPGLFKVQRMNNQEV 60
QY 61 DTVIMFGSYGDRRAVAVKRLMGAQVYKYSKIIIPAAVKIKARDLLIAGMIDTGYFGNT 120
DB 61 DTVIMFGSYGDRRAVAVKRLMGAQVYKYSKIIIPAAVKIKARDLLIAGMIDTGYFGNT 120
QY 121 RVSGIKFIQEDYKQVVDATSVSQIGADTVMNSLGVDGSGVVAIVDTGIDANHPDLK 180
DB 121 RVSGIKFIQEDYKQVVDATSVSQIGADTVMNSLGVDGSGVVAIVDTGIDANHPDLK 180
QY 181 VIGMYDAVNGRSTPYDDQGHGTHVAGIVAGTGSVNSQYIGVAPGAKLVGVKVLGADGSGS 240
DB 181 VIGMYDAVNGRSTPYDDQGHGTHVAGIVAGTGSVNSQYIGVAPGAKLVGVKVLGADGSGS 240
QY 241 VSTIIAGVDVWVQNKDKYGRVNLISGSSQSSDGTDSLQAVNNAMDAGIIVCVAAGNS 300
DB 241 VSTIIAGVDVWVQNKDKYGRVNLISGSSQSSDGTDSLQAVNNAMDAGIIVCVAAGNS 300
QY 301 GPNTYVGSPPAAASKVITVGAVDSDNINAFSSRGPTADRLKPEVVA PGVDIIAPRAGS 360
DB 301 GPNTYVGSPPAAASKVITVGAVDSDNINAFSSRGPTADRLKPEVVA PGVDIIAPRAGS 360
QY 361 TSMGTPINDYTTKASGTSMATPHVSGVGLIIQAHPSMTDPKVTALIFRADIAPRKEIA 420
DB 361 TSMGTPINDYTTKASGTSMATPHVSGVGLIIQAHPSMTDPKVTALIFRADIAPRKEIA 420
QY 421 DIAYGARNVNVYKAIKYDDYAKLFTFGSVADKGSATHTFVSQATFTATILYMDTGSDDI 480
DB 421 DIAYGARNVNVYKAIKYDDYAKLFTFGSVADKGSATHTFVSQATFTATILYMDTGSDDI 480
QY 481 DLVLYDPNGNEVDYSTAYYGFKEVGYNPTAGTWTYKVVSYKGAANYQVDVSDGSLSQ 540
DB 481 DLVLYDPNGNEVDYSTAYYGFKEVGYNPTAGTWTYKVVSYKGAANYQVDVSDGSLSQ 540
QY 541 SGGNNPNPNPNPPTPTDTQTFTGSVNDYMDTSDFTMNVNSGATKLTGDLTFDTSYND 600
DB 541 SGGNNPNPNPNPPTPTDTQTFTGSVNDYMDTSDFTMNVNSGATKLTGDLTFDTSYND 600
QY 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFLLVYASTYGMADYQLKAVVYVG 659
DB 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFLLVYASTYGMADYQLKAVVYVG 659
```

RESULT 3

US-10-090-624-12
Sequence 12, Application US/10090624
Patent No. 6783970
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOTO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshi
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/10/090,624
CURRENT FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: 09/445,472
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patentin version 3.0
SEQ ID NO 12

No BLAST part
problem

LENGTH: 659
TYPE: PRT
ORGANISM: Thermococcus celer
US-10-090-624-12

Query Match 100.0%; Score 3437; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 1.9e-236;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRLGAVNLALVVLGAGTALAAPKPVRRNNAVOOKNYGLTTPGLFKVQRMNNQEV 60
DB 1 MKRLGAVNLALVVLGAGTALAAPKPVRRNNAVOOKNYGLTTPGLFKVQRMNNQEV 60
QY 61 DTVMFGSYGDRRAKVLRLMGAQVKYSYKIIIPAAVAKIKARDLLILGMIDTGYFGNT 120
DB 61 DTVMFGSYGDRRAKVLRLMGAQVKYSYKIIIPAAVAKIKARDLLILGMIDTGYFGNT 120
QY 121 RVSGIKFIOEDYKQVDDATSVSQTADTVVNSLSIGDSGVVAIVDTGIDANHPDLKXK 180
DB 121 RVSGIKFIOEDYKQVDDATSVSQTADTVVNSLSIGDSGVVAIVDTGIDANHPDLKXK 180
QY 181 VIGWTDVAVNGRSTPYDDQGHGTHVAGIVAGTGSVNSQYIGVAPAKLVGVKVLGADGSGS 240
DB 181 VIGWTDVAVNGRSTPYDDQGHGTHVAGIVAGTGSVNSQYIGVAPAKLVGVKVLGADGSGS 240
QY 241 VSTIIAGVDMVYVQNKDKYIRVNLISLSSQSSDGTDSLQAVNNAMDAGIIVCVAAGNS 300
DB 241 VSTIIAGVDMVYVQNKDKYIRVNLISLSSQSSDGTDSLQAVNNAMDAGIIVCVAAGNS 300
QY 301 GNTTVYVSPAAASKYITVGAVDSDNINLASFSSRGPADGRKPEVAVAGVDIIAPRASG 360
DB 301 GNTTVYVSPAAASKYITVGAVDSDNINLASFSSRGPADGRKPEVAVAGVDIIAPRASG 360
QY 361 TSMGTPINDYTKASGSMATPHVSGVALLIQAHSPWPDVKYKALIEADIVAPKEJA 420
DB 361 TSMGTPINDYTKASGSMATPHVSGVALLIQAHSPWPDVKYKALIEADIVAPKEJA 420
QY 421 DIAYAGRVNVYKAIKYDDYAKLFTGSVADKSAHFTFVDSGATFVATLTYMDTSSDI 480
DB 421 DIAYAGRVNVYKAIKYDDYAKLFTGSVADKSAHFTFVDSGATFVATLTYMDTSSDI 480
QY 481 DLYLDPNGNEVDYSTAYYGEKVGYNPTAGTWTVKVYSYKGAANYQVDVSDGSLQ 540
DB 481 DLYLDPNGNEVDYSTAYYGEKVGYNPTAGTWTVKVYSYKGAANYQVDVSDGSLQ 540
QY 541 SGGGNPNPNPNPTPTDTOTFTGSVNDYSDTFTFVNVNSGATKITGDLTFDTSYND 600
DB 541 SGGGNPNPNPNPTPTDTOTFTGSVNDYSDTFTFVNVNSGATKITGDLTFDTSYND 600
QY 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPAGTWTFLVYASTYGMADYOLKAVVY 659
DB 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPAGTWTFLVYASTYGMADYOLKAVVY 659

RESULT 4

US-09-841-553-1
Sequence 1, Application US/09841553
Patent No. 6849441

GENERAL INFORMATION:

APPLICANT: TAKAKURA, Hikaru
MORISHITA, Mio
YAMAMOTO, Katsuhiko
MURTA, Masanori
ASADA, Kiyozo
TSUNASAWA, Susumu
KATO, Ikumoshin
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.

Div. of Biol. Sci.
DBI pat
Biology

COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/841,553
FILING DATE: 24-Apr-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/894,818
FILING DATE: <Unknown>
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-841-553-1

Query Match 100.0%; Score 3437; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 1.9e-236;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRLGAVNLALVVLGAGTALAAPKPVRRNNAVOOKNYGLTTPGLFKVQRMNNQEV 60
DB 1 MKRLGAVNLALVVLGAGTALAAPKPVRRNNAVOOKNYGLTTPGLFKVQRMNNQEV 60
QY 61 DTVMFGSYGDRRAKVLRLMGAQVKYSYKIIIPAAVAKIKARDLLILGMIDTGYFGNT 120
DB 61 DTVMFGSYGDRRAKVLRLMGAQVKYSYKIIIPAAVAKIKARDLLILGMIDTGYFGNT 120
QY 121 RVSGIKFIOEDYKQVDDATSVSQTADTVVNSLSIGDSGVVAIVDTGIDANHPDLKXK 180
DB 121 RVSGIKFIOEDYKQVDDATSVSQTADTVVNSLSIGDSGVVAIVDTGIDANHPDLKXK 180
QY 181 VIGWTDVAVNGRSTPYDDQGHGTHVAGIVAGTGSVNSQYIGVAPAKLVGVKVLGADGSGS 240
DB 181 VIGWTDVAVNGRSTPYDDQGHGTHVAGIVAGTGSVNSQYIGVAPAKLVGVKVLGADGSGS 240
QY 241 VSTIIAGVDMVYVQNKDKYIRVNLISLSSQSSDGTDSLQAVNNAMDAGIIVCVAAGNS 300
DB 241 VSTIIAGVDMVYVQNKDKYIRVNLISLSSQSSDGTDSLQAVNNAMDAGIIVCVAAGNS 300
QY 301 GNTTVYVSPAAASKYITVGAVDSDNINLASFSSRGPADGRKPEVAVAGVDIIAPRASG 360
DB 301 GNTTVYVSPAAASKYITVGAVDSDNINLASFSSRGPADGRKPEVAVAGVDIIAPRASG 360
QY 361 TSMGTPINDYTKASGSMATPHVSGVALLIQAHSPWPDVKYKALIEADIVAPKEJA 420
DB 361 TSMGTPINDYTKASGSMATPHVSGVALLIQAHSPWPDVKYKALIEADIVAPKEJA 420
QY 421 DIAYAGRVNVYKAIKYDDYAKLFTGSVADKSAHFTFVDSGATFVATLTYMDTSSDI 480
DB 421 DIAYAGRVNVYKAIKYDDYAKLFTGSVADKSAHFTFVDSGATFVATLTYMDTSSDI 480
QY 481 DLYLDPNGNEVDYSTAYYGEKVGYNPTAGTWTVKVYSYKGAANYQVDVSDGSLQ 540
DB 481 DLYLDPNGNEVDYSTAYYGEKVGYNPTAGTWTVKVYSYKGAANYQVDVSDGSLQ 540
QY 541 SGGGNPNPNPNPTPTDTOTFTGSVNDYSDTFTFVNVNSGATKITGDLTFDTSYND 600

Db 541 SGGGNPNPNPPTPTDPTQTFGSGVNDWDTSDPTMNVNSGATKITDGLTFDTSYND 600
Qy 601 LDIYLDVDPNGNLVDRSTSSNSYEHEVEYANPAPGWTFLVYASTYGMADYQLAVVYVG 659
Db 601 LDIYLDVDPNGNLVDRSTSSNSYEHEVEYANPAPGWTFLVYASTYGMADYQLAVVYVG 659

RESULT 5
US-08-894-818B-5
; Sequence 5: Application US/08894818B
; Patent No. 6261822
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MITTA, Masanori
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikumoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,818B
; FILING DATE: 20-MAY-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03253
; FILING DATE: 07-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 323285/1995
; FILING DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: TAKAKURA-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-894-818B-5

Query Match 84.8%; Score 2914; DB 2; Length 659;
Best Local Similarity 83.7%; Pred. No. 3.2e-199;
Matches 558; Conservative 47; Mismatches 46; Indels 16; Gaps 7;

Qy 1 MKRLGAVVTLALVGLAGTLLAAPVTPV--VRNNNAVQONVGLITGLFRQVQRMMNQ 58
Db 1 MKGKALITLIVLGLVGVAAAPBKVEQVRN---VERKYGILTPGLFRKIOKLNPNE 57
Qy 59 EVPTVIMFGSVGRDRAVVKLRILMGAVKTSYKIIIPAVAKIKARDLLILAGMDTGVFG 118
Db 58 EISTVIVFENHREKEIIVALELPGAKVRVYTHIIPALADLKLIVISGL--TS--G 113
Qy 119 NTRVSGIKFTQEDYKQVDDA-----TSVSGIGADTVWNSLGYDGSVVVAIVDTGIDAN 173

Db 114 KATLSGVRFTQEDYKQVTSAELEGLDESSAQVATVYVN--LGVDGSGITITGIDTGDAS 172
Qy 174 HPLDKGVKIVGWDVAVNGRSTPYDDQGHGVAGIVAGTGSV--NSQYIGVAPGAKLVGVKV 232
Db 173 HPLDQGVKIVGWDVAVNGRSTPYDDQGHGVAGIVAGTGAASNGKIKGAAPGAKLAGIKV 232
Qy 233 LGADGSGSVSTIIAGVDVWVONKRYGIRVINLSLSSQSSDGTDSLQAVNNAMPAGIV 292
Db 233 LGADGSGSISTIIKGVEMAVDNDKRGIKVINLSLSSQSSDGTDSLQAVNNAMPAGIV 292
Qy 293 VCVAAAGNSGNTTSSPAAASRVITVGANDSNDNLSFSSRPTADGRKPEVVAQVD 352
Db 293 VCVAAAGNSGNTTSSPAAASRVITVGANDSNDNLSFSSRPTADGRKPEVVAQVD 352
Qy 353 IIPRASGSMGPIINDYTTKASGTSMATPHVSGVALILQAPSPTPDKVKTALLETAD 412
Db 353 IIPRASGSMGPIINDYTTKASGTSMATPHVSGVALILQAPSPTPDKVKTALLETAD 412
Qy 413 IYAPKEIADIAYGAGRVNYKAIKYDDYAKLFTGSAVDKSGATHTFDVSGATFTATLY 472
Db 413 IYAPKEIADIAYGAGRVNYKAIKYDDYAKLFTGSAVDKSGATHTFDVSGATFTATLY 472
Qy 473 WDTGSSDIDLXYDPNGNEVDISTAYYGEKYGYNPTAGTWTYVYVSYKGAANYQVDV 532
Db 473 WDTGSSDIDLXYDPNGNEVDISTAYYGEKYGYNPTAGTWTYVYVSYKGAANYQVDV 532
Qy 533 VSDGSLSSQSGGNPNPNPNPTPTDPTQTFGSGVNDWDTSDPTMNVNSGATKITDGL 592
Db 533 VSDGSLSSQSGGNPNPNPNPTPTDPTQTFGSGVNDWDTSDPTMNVNSGATKITDGL 592
Qy 593 TPDTSYNDLXYDPNGNLVDRSTSSNSYEHEVEYANPAPGWTFLVYASTYGMADYQL 652
Db 593 TPDTSYNDLXYDPNGNLVDRSTSSNSYEHEVEYANPAPGWTFLVYASTYGMADYQL 652
Qy 653 KAVVYVG 659
Db 653 KAVVYVG 659

RESULT 6
US-09-841-553-5
; Sequence 5: Application US/09841553
; Patent No. 6849441
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MITTA, Masanori
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikumoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,553
; FILING DATE: 24-Apr-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/894,818
; FILING DATE: <Unknown>
; APPLICATION NUMBER: JP 323285/1995

FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 5
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-841-553-5

Query Match 84.8%; Score 2914; DB 2; Length 659;
Best Local Similarity 83.7%; Pred. No. 3,2e-199;
Matches 558; Conservative 47; Mismatches 46; Indels 16; Gaps 7;

QY 1 MRRLGAVLALVVLGLAAGTALAAPKPV--VRNNAVQOKNYGLTPGLFKYQRMNMQ 58
DB 1 MGKLKALILVILVLGLVGSVAAPKRYQVRN---VEKNYGLTPGLFRKIOKLNPNE 57
QY 59 EVDVTVMFSGYGRDRAVVKVLRMGAVKSYKIIIPAVNKKIARDLLLIAGMIDGYRG 118
DB 58 EISTVIVFENHREKELAVRVLMLGAKRVYVHIIPALIAADLKVDLVISGL--TG--G 113
QY 119 NTRVSGIKFIQEDYKQVDDA-----TSVSIQADTWNLSLGYDSGVVAIVDTGIDAN 173
DB 114 KAKLGSVRFIQEDYKTVSALEGLDESAQVWATYVN--LGYDSGIRITIGITDGDAS 172
QY 174 HPDLKRVIGWYDANVGRSTPYDDQGHGTHVAGIVAGTSV--NSQYIGVAPGAKLVGVK 232
DB 173 HPDLQCKVIGWDFVNGRSYPRYDDHGHGTIVASIAAGTGAASNGKYGAPAKLAGIKV 232
QY 233 LGADSGSVSTIIAGVDWVYQNKDKYKIRINISLSSQSSDGTDSLQAVNNAMADAGIV 292
DB 233 LGADSGSVSTIIKGVEMAVDNKDKYKIRINISLSSQSSDGTDSLQAVNNAMADAGIV 292
QY 293 VCVAAAGNSGPNYTVGSPALAAKVIITVGAVDSDNIIASFSSRGPRTADGRKPEVAPGYD 352
DB 293 VCVAAAGNSGPNYTVGSPALAAKVIITVGAVDSDNIIASFSSRGPRTADGRKPEVAPGYD 352
QY 353 IIAPRASGTSMTGPIINDYTTKASGTSMATPHVSGVALLIQAHPSWTPDKVKTALLETAD 412
DB 353 IIAPRASGTSMTGPIINDYTTKASGTSMATPHVSGVALLIQAHPSWTPDKVKTALLETAD 412
QY 413 IYAPKEIADIAYAGRVNYYKAIKYDDYAKLTFTGSVADKSGATHTFDVSGATFTATLY 472
DB 413 IYAPKEIADIAYAGRVNYYKAIKYDDYAKLTFTGSVADKSGATHTFDVSGATFTATLY 472
QY 473 WPDGSSDIDLXYLDPNGENEVDYTYAYFGEKYGYNPNTAGTWTYVNVSGKGAANYOVY 532
DB 473 WPDGSSDIDLXYLDPNGENEVDYTYAYFGEKYGYNPNTAGTWTYVNVSGKGAANYOVY 532
QY 533 VSDGSLISQSGGPNPNPNPPTTDTOTFTGSVNDYMDTSDFTTMNVNSGATKTIGDL 592
DB 533 VSDGSLISQSGGPNPNPNPPTTDTOTFTGSVNDYMDTSDFTTMNVNSGATKTIGDL 592
QY 593 TPTSTINDLDLYLDNENGLVDRSTSSNYSHEVEYANPAPGTWTFVYVASTYGADYQI 652
DB 593 TPTSTINDLDLYLDNENGLVDRSTSSNYSHEVEYANPAPGTWTFVYVASTYGADYQI 652
QY 653 KAVVYTG 659
DB 653 KAVVYTG 659

RESULT 7
US-08-894-818B-35

Sequence 35, Application US/08894818B
Patent No. 6261822
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MURTA, Masamori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
APPLICANT: KATO, Ikunobu
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSER: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,818B
FILING DATE: 20-MAY-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03253
FILING DATE: 07-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-894-818B-35

Query Match 72.3%; Score 2483.5; DB 2; Length 654;
Best Local Similarity 72.1%; Pred. No. 1.4e-168;
Matches 481; Conservative 69; Mismatches 96; Indels 21; Gaps 8;

QY 1 MRRLGAVLALVVLGLAAGTALAAPKPV--VRNNAVQOKNYGLTPGLFKYQRMNMQ 58
DB 1 MGKLKALILVILVLGLVGSVAAPKRYQVRN---VEKNYGLTPGLFRKIOKLNPNE 57
QY 59 EVDVTVMFSGYGRDRAVVKVLRMGAVKSYKIIIPAVNKKIARDLLLIAGMIDGYRG 118
DB 58 EISTVIVFENHREKELAVRVLMLGAKRVYVHIIPALIAADLKVDLVISGL--TG--G 113
QY 119 NTRVSGIKFIQEDYKQVDDA-----TSVSIQADTWNLSLGYDSGVVAIVDTGIDAN 173
DB 114 KAKLGSVRFIQEDYKTVSALEGLDESAQVWATYVN--LGYDSGIRITIGITDGDAS 172
QY 174 HPDLKRVIGWYDANVGRSTPYDDQGHGTHVAGIVAGTSV--NSQYIGVAPGAKLVGVK 232
DB 173 HPDLQCKVIGWDFVNGRSYPRYDDHGHGTIVASIAAGTGAASNGKYGAPAKLAGIRV 232
QY 233 LGADSGSVSTIIAGVDWVYQNKDKYKIRINISLSSQSSDGTDSLQAVNNAMADAGIV 292
DB 233 LGADSGSVSTIIKGVEMAVDNKDKYKIRINISLSSQSSDGTDSLQAVNNAMADAGIV 292

| | | | | | | | | |
|-----|-----|-----------------|------------------|-------------|---------------|--------------|-----------------|-----|
| Qy | 293 | CVNAGNSGPNYYTYS | GSPALAAKVTITVGA | VDSNDIN | IASFSSSGPTADG | LKXPEVPA | PGVD | 352 |
| Db | 293 | VVVAAGNSGPNKTYT | IGSPALAAKVLITVGA | VDKDYLT | TSFSSSRPTADG | RKXPEVPA | PGVN | 352 |
| Qy | 353 | IIPASGTSWGT | PINDYYTTKASGTS | MAPHYSGVALL | LOAPSMTP | PDKVKAL | ETAD | 412 |
| Db | 353 | IIPAAASGTSWGP | INDYYTTAAPGTSMA | TFPHVAGIAL | LLOAPSMTP | PDKVKAL | ETAD | 412 |
| Qy | 413 | IIVAPKEIADIAG | AGRVNYYKAIKYD | YAKLFTTGS | VADKGSATHTFD | VSAGATFTATLY | 472 | |
| Db | 413 | IYKPEIADIAG | AGRVNYYKAIINY | NYYAKLVFTG | IVANKGSGOTHO | FVIGSAFVATLY | 472 | |
| .Qy | 473 | WDTGSSDIDL | LYYDPNGNEVD | SYTYAYGPEK | GYNNPAGTMT | VVSVKGAANT | QVDY | 532 |
| Db | 473 | MDNANSDIDL | LYYDPNGNOVD | SYTAYYGEK | GYNNPFDGWT | IKVBSIGSANT | QVDY | 532 |
| Qy | 533 | VSDGSLSGGGG | GNPNPNPNPPT | PTDTOTFTGS | VNDYMTSDT | FTMNVS | GATKITGDL | 592 |
| Db | 533 | VSDGSLSG---- | PGSSPSBPQEP | PTDAKITFG | SDSHYYIDR | STFTMTVS | GATKITGDL | 592 |
| Qy | 593 | TEDTSTYNDL | LYLYDENGM | LVDRSTSSNS | IEHVEYAN | PAEGTMTFL | VAYASTYGMADYOL | 652 |
| Db | 588 | VFDTSTYHDL | LYLYDENQ | LVDRSES | PNSIEHVEY | LTLPAGTMY | FLVAYASTYGMAYEL | 647 |
| Qy | 653 | KAVVYTYG | 659 | | | | | |
| Db | 648 | TAKVYTYG | 654 | | | | | |

RESULT 8
US-09-445-472-16
; Sequence 16, Application US/09445472

```

/ GENERAL INFORMATION:
/ APPLICANT: TAKAKURA, Hikaru
/ APPLICANT: MORISHITA, Mio
/ APPLICANT: SHIMOJO, Tomoko
/ APPLICANT: ASADA, Kiyozo
/ APPLICANT: KATO, Ikuneshin
/ TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
/ FILE REFERENCE: TAKAKURA=6
/ CURRENT APPLICATION NUMBER: US/09/445,472
/ PRIOR FILING DATE: 1999-12-06
/ PRIOR APPLICATION NUMBER: 151969/1997
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 16
/
/ LENGTH: 654
/
/ TYPE: PR1
/ ORGANISM: Pyrococcus furiosus
/ US-09-445,472-16

```

| | | | | |
|---------------------------|--------|---------------------|------------|-------------|
| Query Match | 72.3%; | Score 2483.5; | DB 2; | Length 654; |
| Best Local Similarity | 72.1%; | Pred. No. 1.4e-168; | | |
| Matches 481; Conservative | 63; | Mismatches 96; | Indels 21; | Gaps 8; |

```

Oy 1 MRLRLAVALVATLVYGLAGTALAAAPKPV--VANNNAVOOKNGVLLPGLPKKYQÖMMNNO 58
Db 1 MGGLKALLVITLVVLGVGSDVAAPAKPKKQVÖVN---VEKNQGLLTPGLRPRKKIÖQKNPNP 57
Oy 59 EYDVTVMFPGSGYGRDRAVKVLRMLMGQOVKSYKIIPAVAVKIIARLDLLTAGMIDTGYFG 118
Db 58 EISTYVTFENHREKELAAKVLELMGAKKRVYHIIIRIADLDKVRDLLVTSGL--TG--G 113
Oy 119 NTRVSGIKETIOBDYKQVÖDDA-----TSVSQIGADTVMNSLGYDGSQVVAIVDTGIDAN 173
Db 114 KAKLSGVRFIQEDBYKTVTSABTEGLDEBSAAQVAAVATVMN--LGYDGSQITIGIIDTGIDAS 172
Oy 174 HPDLKCKVTGWDVAVNGRSTPYDDQGHGHTVAGIVAGTGSV--NSQYIGVAPGAKLVGVKV 232
Db 173 HEDLOQKVIQWDFVNGRSTPYDDHGHGHTVASIAGTGAASNGKYGKAPGAKLGIKV 232

```

| | | | |
|----|-----|---|-----|
| QY | 233 | GAAGSGSVSTIIAGVDMVYVONKDYKGRVNLSTGSSQSSPBGTSLSQAVNNAMDAQIV | 292 |
| Db | 233 | LGADSGSISITIIKGVEMAVDNKDKYKJLVINLSIGSSQSSSDGTDALSQAIVAMDAQIV | 292 |
| QY | 293 | VCVAAGNSGPNYTTVGSPPAASKVITTVGAVDSDNDIASFSRSGPDAADGLKCPVYAPGV | 352 |
| Db | 293 | VYVAAGNSGPNKYRTIGSPPAASKVITTVGAVDYKDYVITFSRSGPDAADGLKCPVYAPGV | 352 |
| QY | 353 | IIAPRAGTSMGTPINDYYTKASGTSMAIIPHVSGALIILOHPSTPCKVXTALLETAD | 412 |
| Db | 353 | IIAASASTSMGQPLNDYYTAAEGTSMAIIPHVAGIALLOHPSTPCKVXTALLETAD | 412 |
| QY | 413 | IVAPKEIADIAYGAGRNVVYKAIKIDYDPAKLFTGTSVADKSGATHFDVSGATFVATLTY | 472 |
| Db | 413 | IYKPEIADIAYGAGRNVAYKAIKINDYDAKLFTGTAVANKSGQTHQFVIGASFTVATLTY | 472 |
| QY | 473 | WDTGSSDIDLYLYDPNGNEVDYSTYAYYGFERYGYNPTPAGTMYTVYVYSKCAANYQVUD | 532 |
| Db | 473 | WDMNANSDDLYLYDPNGNQVDYSYAYYGFERYGYNPTPDGTMITKVYSYSSANQVUD | 532 |
| QY | 533 | VSDGSLSGGGGNPNPNPNPTPTDQTFGGSVNDYWDTSIDTFTMANVSGATKITDGL | 592 |
| Db | 533 | VSDGSLSGQ----PSSPSPOPEPYYDAKTFGGSDBHYIYDRSDITFTMYNSGATKITDGL | 587 |
| QY | 593 | TFDTSYNDLJLYLYDPNGNLVDRSTSSNSEYHEVYANPAEGTWTFLVYAYSTYGMADYOL | 652 |
| Db | 588 | VFDTSYHDLJLYLYDPNQKLVDRSSPNSYHEVEYLTLPAGTMYFLVYAYTYGMAYYEL | 647 |
| QY | 653 | KAVYTYG 659 | |
| Db | 648 | TAKYTYG 654 | |

RESULT 9
US-10-090-624-16
; Sequence 16, Application US/10090624

```

APPLICANT : TAKAKURA, Hikaru
APPLICANT : MORISHITA, Mio
APPLICANT : SHIMOJO, Tomoko
APPLICANT : ASADA, Kiyozo
APPLICANT : KATO, Ikumoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/10/090,624
CURRENT FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: 08/445,472
PRIORITY FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIORITY FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 16
LENGTH: 654
TYPE: PRT
ORGANISM: Pyrococcus furiosus
US-10-090-624-16
```

| | | | |
|---------------------------|--------|---------------------|--------------------|
| Query Match | 72.3%; | Score 2483.5; | Length 654; |
| Best Local Similarity | 72.1%; | Pred. No. 1.4e-168; | |
| Matches 481; Conservative | 69; | Mismatches 96; | Indels 21; Gaps 8; |

[illegible]

```

Db      114 KAKLSGVRFOEDKYKTVSAELBGLDESAQVWATYVMN-LGVDGSGITIGITDGDAS 172
      174 HPDLKGVIGWVAVNVRSTPYDDOGHGTHTVAGI VAGTGSV-NSQYGVAPGAKLVGVKY 232
      173 HPDLQGVIGWVAVNVRSTPYDDHGHGTHTVAGI VAGTGSAGKYGKAPGAKLVGIRY 232
      233 LGADGSGSVSTIAGVDMVQNKDKYGI RIVINLSLSSGSSSDGDTSLQAVNNAMPAGIY 232
      233 LGADGSGSVSTIAGVDMVQNKDKYGI RIVINLSLSSGSSSDGDTSLQAVNNAMPAGIY 232
      293 VCVAAAGNSGPNYTVGSPAAASKVITVGA VDSNDNIASFSSRGPADGRKPEVAVPGVD 352
      293 VCVAAAGNSGPNYTVGSPAAASKVITVGA VDSNDNIASFSSRGPADGRKPEVAVPGVD 352
      353 IIAPRASGTMGPINDYTYKASGTSMATPHVSGVALLIQAHPSWTPDKVKTALJETAD 412
      353 IIAPRASGTMGPINDYTYKASGTSMATPHVSGVALLIQAHPSWTPDKVKTALJETAD 412
      413 IYAPKEIADIAYAGAGRVNVAIKKYDDYAKLFTGSAVADKGSATHTFVDSGATEFVATLY 472
      413 IYAPKEIADIAYAGAGRVNVAIKKYDDYAKLFTGSAVADKGSATHTFVDSGATEFVATLY 472
      473 WDTGSSDIDLXYDPNGNEVDYSYTAAYGFEKYGYNPTAGTWTVKVVSYKGAANYQVDY 532
      473 WDTGSSDIDLXYDPNGNEVDYSYTAAYGFEKYGYNPTAGTWTVKVVSYKGAANYQVDY 532
      533 VSDGSLSSQSGGPNPNPNPTPTDTQTFGTVSVDYWDYDTSPTFMNNSGATKTITGDL 552
      533 VSDGSLSSQSGGPNPNPNPTPTDTQTFGTVSVDYWDYDTSPTFMNNSGATKTITGDL 552
      533 VSDGSLSSQSGGPNPNPNPTPTDTQTFGTVSVDYWDYDTSPTFMNNSGATKTITGDL 552
      593 TPDTSYNDLXYDPNGNLYDRSSTSSNSYEHVEYANPARGTWTFLVYASTYGMADYOL 652
      593 TPDTSYNDLXYDPNGNLYDRSSTSSNSYEHVEYANPARGTWTFLVYASTYGMADYOL 652
      647 VPTSTHDLXYDPNGNLYDRSSTSSNSYEHVEYANPARGTWTFLVYASTYGMADYOL 647
      648 TAKVYTG 654

RESULT 10
US-09-841-553-35
; Sequence 35, Application US/09841553
; Patent No. 6843441
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; MORISHITA, Mio
; YAMAMOTO, Katsuhiko
; MITTA, Masanori
; ASADA, Kiyozo
; TSUNASAWA, Susumu
; KATO, Ikunoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; City: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,553
; FILING DATE: 24-Apr-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/894,818
; FILING DATE: <Unknown>
; APPLICATION NUMBER: JP 323285/1995

```

```

; FILING DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: TAKAKURA-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 654 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-841-553-35

Query Match      72.3%; Score 2483.5; DB 2; Length 654;
Best Local Similarity 72.1%; Pred. No. 1.4e-168;
Matches 481; Conservative 69; Mismatches 96; Indels 21; Gaps 8;

      1 MKRLGAVIALVVLVGLAAGTALAAPYKPV--VRNNAVQOKNTGLTPGLFKYQQRNMQ 58
      1 MKGLKALLIIVILVGLVGSVAAPPEKKEOVARN--VEKNYGLTLPGFLFRKIQKLNPN 57
      59 EVDVTIMFGSYGDRDAVAVVFLMGQVYKSKITIPAVAVKIKARPLLIIAGMIDGYRG 118
      58 EISTVIVFENHREKELAVVLELMGAKRVYVHIIPALADLVKRLVLSGL--TG--G 113
      119 NTRVSGIKFIOEDYKQVQVDA-----TSVQICADPTVMSLGYDGSVVAIVDTGIDAN 173
      114 KAKLSGVRFOEDKYKTVSAELBGLDESAQVWATYVMN-LGVDGSGITIGITDGDAS 172
      174 HPDLKGVIGWVAVNVRSTPYDDOGHGTHTVAGI VAGTGSV-NSQYGVAPGAKLVGVKY 232
      173 HPDLQGVIGWVAVNVRSTPYDDHGHGTHTVAGI VAGTGSAGKYGKAPGAKLVGIRY 232
      233 LGADGSGSVSTIAGVDMVQNKDKYGI RIVINLSLSSGSSSDGDTSLQAVNNAMPAGIY 232
      233 LGADGSGSVSTIAGVDMVQNKDKYGI RIVINLSLSSGSSSDGDTSLQAVNNAMPAGIY 232
      293 VCVAAAGNSGPNYTVGSPAAASKVITVGA VDSNDNIASFSSRGPADGRKPEVAVPGVD 352
      293 VCVAAAGNSGPNYTVGSPAAASKVITVGA VDSNDNIASFSSRGPADGRKPEVAVPGVD 352
      353 IIAPRASGTMGPINDYTYKASGTSMATPHVSGVALLIQAHPSWTPDKVKTALJETAD 412
      353 IIAPRASGTMGPINDYTYKASGTSMATPHVSGVALLIQAHPSWTPDKVKTALJETAD 412
      413 IYAPKEIADIAYAGAGRVNVAIKKYDDYAKLFTGSAVADKGSATHTFVDSGATEFVATLY 472
      413 IYAPKEIADIAYAGAGRVNVAIKKYDDYAKLFTGSAVADKGSATHTFVDSGATEFVATLY 472
      473 WDTGSSDIDLXYDPNGNEVDYSYTAAYGFEKYGYNPTAGTWTVKVVSYKGAANYQVDY 532
      473 WDTGSSDIDLXYDPNGNEVDYSYTAAYGFEKYGYNPTAGTWTVKVVSYKGAANYQVDY 532
      533 VSDGSLSSQSGGPNPNPNPTPTDTQTFGTVSVDYWDYDTSPTFMNNSGATKTITGDL 552
      533 VSDGSLSSQSGGPNPNPNPTPTDTQTFGTVSVDYWDYDTSPTFMNNSGATKTITGDL 552
      593 TPDTSYNDLXYDPNGNLYDRSSTSSNSYEHVEYANPARGTWTFLVYASTYGMADYOL 652
      588 VPTSTHDLXYDPNGNLYDRSSTSSNSYEHVEYANPARGTWTFLVYASTYGMADYOL 647
      653 KAVVYTG 659
      648 TAKVYTG 654

RESULT 11
US-08-894-818B-3

```


Sequence 3, Application US/08894818B
Patent No. 6261822
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MITTA, Masanori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,818B
FILING DATE: 20-MAY-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03253
FILING DATE: 07-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 737-3528
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 522 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: /note=Xaa at position 428 is Gly or Val.
US-08-894-818B-3
Query Match 62.2%; Score 2138.5; DB 2; Length 522;
Best Local Similarity 78.3%; Pred. No. 3.9e-144;
Matches 407; Conservative 36; Mismatches 70; Indels 7; Gaps 3;
QY 141 SVSGIGADTVNWSLIGDSSGVVAIVDTGIDANHPDLKGVIGWYDAVNGRSTPYDDOQH 200
DB 9 SAAQVATYVWN-LGYDSSGITIGIITGIDASHPLDQKVIQWDPVNGRSTPYDDHGH 67
QY 201 GTHVAGIVAGTGSV-NSQYIGVAPGAKLVGVKVLGADSGSVSTIIAGVDMVYQNKDKYXG 259
DB 68 GTHVASIAAGTGAASNGKYGMAPGAKLAGIKVLGADSGSISTIIKGVMAVNDKDKYXG 127
QY 260 IRVINISLSSQSSDGTSLDLSQAVNNAMDAIVVCVAAGNSGPNYTVVGSPPAASKVITV 319
DB 128 IKVINISLSSQSSDGTDLALSQAVNNAMDAIVVCVAAGNSGPNKTTIGSPAAASKVITV 187
QY 320 GAVDSNDNISFSSRGRTADGRLKPEYVAPGVDIIPAPASGSMGTPINDYTTKASGTSM 379
DB 188 GAVDKYDVITFSFSSRGRTADGRLKPEYVAPGVNMIIPAPASGSMGQPINDYTTAAGTSM 247
QY 380 ATPHVSQVALLIQAPHSWTPDKVKTALITETADIVAPKEIADIAYGAGRNVYKAIKYDD 439

DB 248 ATPHVAGIAALLIQAPHSWTPDKVKTALITETADIVAPKEIADIAYGAGRNVYKAIKYDD 307
QY 440 YAKLITFGSVADKGSATHTFVSGATFTVATLTWYDSSSIDLYLDPKNEVDYSYTA 439
DB 308 YAKLVFTGVVANKSGQTHQFVSGATFTVATLTWYDSSSIDLYLDPKNOVDYSYTA 367
QY 500 YGFEKGYVNPITGTVTKVSVYKGAANYQVDVVSQSLSSQSGGNPNPNPPTPTD 559
DB 368 YGFEKGYVNPITGTVTKVSVYSGSANYQVDVVSQSLSSQSGGNPNPNPPTPTD 422
QY 560 TQFTFGSVNDYDTSPTFTMNVNSGATKTTGDTFTPTSYNDLILYLDPNGLVDRSTSS 619
DB 423 AKTFQXSDHYVYDRSDFTFTVNSGATKTTGDLVFTPTSYNDLILYLDPNGLVDRSESP 482
QY 620 NSYEHVEYANPARGTWTFLVVAYSTGMDYOLKAVVYVG 659
DB 483 NSYEHVEYLTAPGVTWFLVVAYSTGMDYOLKAVVYVG 522
RESULT 12
US-09-445-472-4
Sequence 4, Application US/09445472
Patent No. 6358726
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOJO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA=6
CURRENT APPLICATION NUMBER: US/09/445,472
CURRENT FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 522
TYPE: PRT
ORGANISM: Pyrococcus furiosus
FEATURE:
NAME/KEY: misc_feature
LOCATION: (428)..(428)
OTHER INFORMATION: Xaa at position 428 is Gly or Val.
US-09-445-472-4
Query Match 62.2%; Score 2138.5; DB 2; Length 522;
Best Local Similarity 78.3%; Pred. No. 3.9e-144;
Matches 407; Conservative 36; Mismatches 70; Indels 7; Gaps 3;
QY 141 SVSGIGADTVNWSLIGDSSGVVAIVDTGIDANHPDLKGVIGWYDAVNGRSTPYDDOQH 200
DB 9 SAAQVATYVWN-LGYDSSGITIGIITGIDASHPLDQKVIQWDPVNGRSTPYDDHGH 67
QY 201 GTHVAGIVAGTGSV-NSQYIGVAPGAKLVGVKVLGADSGSVSTIIAGVDMVYQNKDKYXG 259
DB 68 GTHVASIAAGTGAASNGKYGMAPGAKLAGIKVLGADSGSISTIIKGVMAVNDKDKYXG 127
QY 260 IRVINISLSSQSSDGTSLDLSQAVNNAMDAIVVCVAAGNSGPNYTVVGSPPAASKVITV 319
DB 128 IKVINISLSSQSSDGTDLALSQAVNNAMDAIVVCVAAGNSGPNKTTIGSPAAASKVITV 187
QY 320 GAVDSNDNISFSSRGRTADGRLKPEYVAPGVDIIPAPASGSMGTPINDYTTKASGTSM 379
DB 188 GAVDKYDVITFSFSSRGRTADGRLKPEYVAPGVNMIIPAPASGSMGQPINDYTTAAGTSM 247
QY 380 ATPHVSQVALLIQAPHSWTPDKVKTALITETADIVAPKEIADIAYGAGRNVYKAIKYDD 439
DB 248 ATPHVAGIAALLIQAPHSWTPDKVKTALITETADIVAPKEIADIAYGAGRNVYKAIKYDD 307
QY 440 YAKLITFGSVADKGSATHTFVSGATFTVATLTWYDSSSIDLYLDPKNEVDYSYTA 439


```
Db      9  SAQVMAATYWN--LGVDGSGITIGI IDTG:IDASHPDLQGVKIVGWVDFVNGRSTPYDDHG 67
Qy      201 GTHVAGI VAGTGSV--NSQYIGVAPGATLVGVKTLGADGSSVSTIIGVWVYQNDKYG 259
Db      68  GTHVASTIAACTGAASNKYKGMAPGATLAIKYLGAAGSSISTIIKGVMAVDNKKYK 127
Qy      260 IRVINLSLSSQSSDGTDSLQA VNNMADAGI VVCAAGNSGPNYTVGSPAASKYITV 319
Db      128 IKVINLSLSSQSSDGTDLALQA VNNMADAGI VVVAAGNSGPNKYITGSPAASKYITV 187
Qy      320 GAVDSNDNIA SFSSRGPTADGRLKPEVVA PGVDIIAPRASGTMGPINDYTTKASGTM 379
Db      188 GAVDKYDVITSFSSRGPTADGRLKPEVVA PGVNI IAPRASGTMGPINDYTTAAGTSM 247
Qy      380 ATPHVSQVGLIIQAHPSWTPDKVKTAL IETADI VAPKEIADIA YGAGRNVYKAIKYDD 439
Db      248 ATPHVAGIALLLQAHPSWTPDKVKTAL IETADIVKPEIADIA YGAGRNVYKAIKYDD 307
Qy      440 YAKLFTGSVADKGSATHTFDVSGATFVATLYMTDSSDIDL YLDPNGNEVDYSYAY 499
Db      308 YALVFTGYVANKSGQTHQFVIGSASFVATLYMDNANSDDL YLDPNGNQVDYSYAY 367
Qy      500 YGEEKYGYNPTAGTWTVKVSYKGAANYQVDVSDGSLSGGQGNPNPNPPTTD 559
Db      368 YGEEKYGYNPTDGTWTIKVSYSGSANYQVDVSDGSLSG-----FGSSSPQPEPTVD 422
Qy      560 TQFTGSVNDYMTDPTDFETMNVNSGATKITGDLTPDTSYNDL DLYLDPNGNIVDRSTG 619
Db      423 AKTFQASDHYTDRSDIFTMTVNSGATKITGDLVFDTSYHDDL YLDPNGQLVDRSESP 482
Qy      620 NSYEHVEYANPAPGTWTLVYAYSTYGMADYOLKAVVYVG 659
Db      483 NSYEHVEYLTLPAPGTWTLVYAYTYGMAYYBLAKVYVG 522

RESULT 15
US-09-445-472-1
; Sequence 1, Application US/09445472
; Patent No. 6358726
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-09-445-472-1
```

```
Query Match      49.7%; Score 1707; DB 2; Length 412;
Best Local Similarity 80.6%; Pred. No. 1,5e-113;
Matches 325; Conservative 29; Mismatches 47; Indels 2; Gaps 2;
```

```
Qy      141 SVSODIADTYWNSLGYDGSQVVAIYDTGIDANHPDLKGVIGVYDAVNGRSTPYDDQH 200
Db      9  SAQVMAATYWN--LGVDGSGITIGI IDTG:IDASHPDLQGVKIVGWVDFVNGRSTPYDDHG 67
Qy      201 GTHVAGI VAGTGSV--NSQYIGVAPGATLVGVKTLGADGSSVSTIIGVWVYQNDKYG 259
Db      68  GTHVASTIAACTGAASNKYKGMAPGATLAIKYLGAAGSSISTIIKGVMAVDNKKYK 127
Qy      260 IRVINLSLSSQSSDGTDSLQA VNNMADAGI VVCAAGNSGPNYTVGSPAASKYITV 319
```

```
Db      128 IKVINLSLSSQSSDGTDLALQA VNNMADAGI VVVAAGNSGPNKYITGSPAASKYITV 187
Qy      320 GAVDSNDNIA SFSSRGPTADGRLKPEVVA PGVDIIAPRASGTMGPINDYTTKASGTM 379
Db      188 GAVDKYDVITSFSSRGPTADGRLKPEVVA PGVNI IAPRASGTMGPINDYTTAAGTSM 247
Qy      380 ATPHVSQVGLIIQAHPSWTPDKVKTAL IETADI VAPKEIADIA YGAGRNVYKAIKYDD 439
Db      248 ATPHVAGIALLLQAHPSWTPDKVKTAL IETADIVKPEIADIA YGAGRNVYKAIKYDD 307
Qy      440 YAKLFTGSVADKGSATHTFDVSGATFVATLYMTDSSDIDL YLDPNGNEVDYSYAY 499
Db      308 YALVFTGYVANKSGQTHQFVIGSASFVATLYMDNANSDDL YLDPNGNQVDYSYAY 367
Qy      500 YGEEKYGYNPTAGTWTVKVSYKGAANYQVDVSDGSLSGG 542
Db      368 YGEEKYGYNPTDGTWTIKVSYSGSANYQVDVSDGSLSG 410
```

```
Search completed: January 6, 2007, 22:33:39
Job time : 37 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2007 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: January 6, 2007, 22:31:42 ; Search time 36 Seconds
(without alignments)
1602.297 Million cell updates/sec

Title: US-10-800-684-5

Perfect score: 3428
Sequence: 1 MKGLKALILVILVGLHVG.....YASTYGMADYOLKAVVYVG 659

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued Patents NA:
2: /EMC_Celerra_SIDS3/Pctodata/2/1aa/5/COMB.pep:*
3: /EMC_Celerra_SIDS3/Pctodata/2/1aa/6/COMB.pep:*
4: /EMC_Celerra_SIDS3/Pctodata/2/1aa/7/COMB.pep:*
5: /EMC_Celerra_SIDS3/Pctodata/2/1aa/H/COMB.pep:*
6: /EMC_Celerra_SIDS3/Pctodata/2/1aa/PCTUS.COMB.pep:*
7: /EMC_Celerra_SIDS3/Pctodata/2/1aa/RE.COMB.pep:*
8: /EMC_Celerra_SIDS3/Pctodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|--------|-------------|-----------|----|-------------------|
| 1 | 3428 | 100.0 | 659 | 2 | US-08-894-818B-5 |
| 2 | 3428 | 100.0 | 659 | 2 | US-09-841-553-5 |
| 3 | 2997.5 | 87.4 | 654 | 2 | US-08-894-818B-35 |
| 4 | 2997.5 | 87.4 | 654 | 2 | US-09-445-472-16 |
| 5 | 2997.5 | 87.4 | 654 | 2 | US-10-090-624-16 |
| 6 | 2997.5 | 87.4 | 654 | 2 | US-09-841-553-35 |
| 7 | 2997.5 | 87.4 | 659 | 2 | US-08-894-818B-1 |
| 8 | 2997.5 | 87.4 | 659 | 2 | US-09-445-472-12 |
| 9 | 2997.5 | 87.4 | 659 | 2 | US-10-090-624-12 |
| 10 | 2997.5 | 87.4 | 659 | 2 | US-09-841-553-1 |
| 11 | 2997.5 | 87.4 | 652 | 2 | US-08-894-818B-3 |
| 12 | 2997.5 | 87.4 | 652 | 2 | US-09-445-472-4 |
| 13 | 2997.5 | 87.4 | 652 | 2 | US-10-090-624-4 |
| 14 | 2997.5 | 87.4 | 652 | 2 | US-09-841-553-3 |
| 15 | 2997.5 | 87.4 | 652 | 2 | US-09-445-472-1 |
| 16 | 2997.5 | 87.4 | 652 | 2 | US-10-090-624-1 |
| 17 | 2997.5 | 87.4 | 652 | 2 | US-08-894-818B-18 |
| 18 | 2997.5 | 87.4 | 734 | 2 | US-09-000-016-4 |
| 19 | 2997.5 | 87.4 | 734 | 2 | US-09-514-340-4 |
| 20 | 2997.5 | 87.4 | 823 | 2 | US-09-000-016-2 |
| 21 | 2997.5 | 87.4 | 823 | 2 | US-09-514-340-2 |
| 22 | 2997.5 | 87.4 | 520 | 2 | US-09-000-016-7 |
| 23 | 2997.5 | 87.4 | 520 | 2 | US-09-514-340-7 |
| 24 | 2997.5 | 87.4 | 329 | 1 | US-08-885-774-2 |
| 25 | 2997.5 | 87.4 | 382 | 2 | US-09-445-472-31 |
| 26 | 2997.5 | 87.4 | 382 | 2 | US-10-090-624-31 |

| | | | | | | |
|----|-------|------|-----|---|------------------|--------------------|
| 27 | 555 | 16.2 | 382 | 2 | US-10-104-693-2 | Sequence 2, Appl1 |
| 28 | 554.5 | 16.2 | 378 | 1 | US-07-772-087-4 | Sequence 4, Appl1 |
| 29 | 554 | 16.2 | 379 | 1 | US-08-845-295A-1 | Sequence 1, Appl1 |
| 30 | 554 | 16.2 | 379 | 2 | US-09-140-933-1 | Sequence 1, Appl1 |
| 31 | 554 | 16.2 | 379 | 2 | US-09-146-661-1 | Sequence 1, Appl1 |
| 32 | 554 | 16.2 | 379 | 2 | US-09-150-515-1 | Sequence 1, Appl1 |
| 33 | 551 | 16.1 | 382 | 2 | US-09-355-502-2 | Sequence 2, Appl1 |
| 34 | 551 | 16.1 | 382 | 2 | US-09-659-749-2 | Sequence 2, Appl1 |
| 35 | 551 | 16.1 | 382 | 2 | US-09-672-105-2 | Sequence 2, Appl1 |
| 36 | 550 | 16.0 | 382 | 7 | 5472855-2 | Patent No. 5472855 |
| 37 | 549.5 | 16.0 | 381 | 1 | US-07-772-087-6 | Sequence 6, Appl1 |
| 38 | 549.5 | 16.0 | 381 | 1 | US-08-173-508-12 | Sequence 12, Appl1 |
| 39 | 549.5 | 16.0 | 381 | 1 | US-08-265-310-12 | Sequence 12, Appl1 |
| 40 | 549.5 | 16.0 | 381 | 2 | US-08-951-742-12 | Sequence 2, Appl1 |
| 41 | 547 | 16.0 | 382 | 2 | US-09-178-155-2 | Sequence 2, Appl1 |
| 42 | 547 | 16.0 | 382 | 2 | US-09-178-173A-2 | Sequence 2, Appl1 |
| 43 | 547 | 16.0 | 382 | 2 | US-09-177-353-2 | Sequence 2, Appl1 |
| 44 | 547 | 16.0 | 382 | 2 | US-10-033-325-2 | Sequence 2, Appl1 |
| 45 | 547 | 16.0 | 382 | 2 | US-10-828-572-2 | Sequence 2, Appl1 |

ALIGNMENTS

RESULT 1
US-08-894-818B-5
Sequence 5, Application US/08894818B
Patent No. 6261822
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MITTA, Masamori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
APPLICANT: KATO, Ikumoshin
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSER: Broadway and Nelmark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894, 818B
CLASSIFICATION: 435
FILING DATE: 20-May-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03253
FILING DATE: 07-Nov-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

*Div. of this infant
reference
Main Seq id no. 3 (p.p)
No Db.pat. problems*

MOLECULE TYPE: peptide
US-08-894-818B-5

Query Match 100.0%; Score 3428; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 3.9e-241;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKGKALILVILVGLVGVSAAPKPKVQVNEKNYGLTPGLFRKIQKLNPEEIS 60
DB 1 MKGKALILVILVGLVGVSAAPKPKVQVNEKNYGLTPGLFRKIQKLNPEEIS 60
QY 61 TVIVFENHREKEIAVRLELMGAKRVYVHIIPALADLVKVDLVIISGLTGKAKLSGV 120
DB 61 TVIVFENHREKEIAVRLELMGAKRVYVHIIPALADLVKVDLVIISGLTGKAKLSGV 120
QY 121 RFIQEDYKTVSAELFGLDESAQVMAVYVNLGYDGSITIGITIDGIDASHPDLOGKV 180
DB 121 RFIQEDYKTVSAELFGLDESAQVMAVYVNLGYDGSITIGITIDGIDASHPDLOGKV 180
QY 181 IGWVDFVNGRSYPYDDHGHGTHVASIAAGTGAASNGKYGMAPGAAGIIVCVAAAGS 240
DB 181 IGWVDFVNGRSYPYDDHGHGTHVASIAAGTGAASNGKYGMAPGAAGIIVCVAAAGS 240
QY 241 ISTTIKGVEMAVNDKDKYGIKVINLSGSSQSSDGTSLSOAVNNAMDAGIIVCVAAAGS 300
DB 241 ISTTIKGVEMAVNDKDKYGIKVINLSGSSQSSDGTSLSOAVNNAMDAGIIVCVAAAGS 300
QY 301 GPTTYVGSPPAASKVTITVGAVDNDNIASFSSRGPADGRLEKPEVAVPGVDIIPRAGS 360
DB 301 GPTTYVGSPPAASKVTITVGAVDNDNIASFSSRGPADGRLEKPEVAVPGVDIIPRAGS 360
QY 361 TSMGTPINDYTKASGSMATPHVSGVALLIQAHPSTPDKXTALIFRADIVAPKEIA 420
DB 361 TSMGTPINDYTKASGSMATPHVSGVALLIQAHPSTPDKXTALIFRADIVAPKEIA 420
QY 421 DIAYGAGRNVVYRAIKYDDYAKLFTFGSVADKGSATTFPVSGATFVTATLYMDTGSDDI 480
DB 421 DIAYGAGRNVVYRAIKYDDYAKLFTFGSVADKGSATTFPVSGATFVTATLYMDTGSDDI 480
QY 481 DLYLDPNGNEVDYSYTAAYGFEKVGYNPTAGTWTVKVSYKGAANYQVDVSDGSLSQ 540
DB 481 DLYLDPNGNEVDYSYTAAYGFEKVGYNPTAGTWTVKVSYKGAANYQVDVSDGSLSQ 540
QY 541 SGGGNPNPNPNPTPTDPTFTGTVNDYMDTFTFMANVSGATKITGDLTFDTSYND 600
DB 541 SGGGNPNPNPNPTPTDPTFTGTVNDYMDTFTFMANVSGATKITGDLTFDTSYND 600
QY 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPARGTWTFVYVASTYGADYQLKAVYYG 659
DB 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPARGTWTFVYVASTYGADYQLKAVYYG 659

RESULT 2
US-09-841-553-5

Sequence 5, Application US/09841553
Patent No. 6849441

GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru

MORISHITA, Mio
YAMAMOTO, Katsuhiko

MITTA, Masanori
ASADA, Kiyozo

TSUNASAWA, Susumu
KATO, Ikunoshin

TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:
ADDRESSES: Browdy and Neimark
CITY: Washington

STATE: D. C.
COUNTRY: United States of America

ZIP: 20004

Handwritten notes:
"of interest" (circled)
"only dist. 5-7 id no. 152."
"No. 152 not found."

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/841,553
FILING DATE: 24-Apr-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/894,818
FILING DATE: <Unknown>
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 5
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-841-553-5

Query Match 100.0%; Score 3428; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 3.9e-241;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKGKALILVILVGLVGVSAAPKPKVQVNEKNYGLTPGLFRKIQKLNPEEIS 60
DB 1 MKGKALILVILVGLVGVSAAPKPKVQVNEKNYGLTPGLFRKIQKLNPEEIS 60
QY 61 TVIVFENHREKEIAVRLELMGAKRVYVHIIPALADLVKVDLVIISGLTGKAKLSGV 120
DB 61 TVIVFENHREKEIAVRLELMGAKRVYVHIIPALADLVKVDLVIISGLTGKAKLSGV 120
QY 121 RFIQEDYKTVSAELFGLDESAQVMAVYVNLGYDGSITIGITIDGIDASHPDLOGKV 180
DB 121 RFIQEDYKTVSAELFGLDESAQVMAVYVNLGYDGSITIGITIDGIDASHPDLOGKV 180
QY 181 IGWVDFVNGRSYPYDDHGHGTHVASIAAGTGAASNGKYGMAPGAAGIIVCVAAAGS 240
DB 181 IGWVDFVNGRSYPYDDHGHGTHVASIAAGTGAASNGKYGMAPGAAGIIVCVAAAGS 240
QY 241 ISTTIKGVEMAVNDKDKYGIKVINLSGSSQSSDGTSLSOAVNNAMDAGIIVCVAAAGS 300
DB 241 ISTTIKGVEMAVNDKDKYGIKVINLSGSSQSSDGTSLSOAVNNAMDAGIIVCVAAAGS 300
QY 301 GPTTYVGSPPAASKVTITVGAVDNDNIASFSSRGPADGRLEKPEVAVPGVDIIPRAGS 360
DB 301 GPTTYVGSPPAASKVTITVGAVDNDNIASFSSRGPADGRLEKPEVAVPGVDIIPRAGS 360
QY 361 TSMGTPINDYTKASGSMATPHVSGVALLIQAHPSTPDKXTALIFRADIVAPKEIA 420
DB 361 TSMGTPINDYTKASGSMATPHVSGVALLIQAHPSTPDKXTALIFRADIVAPKEIA 420
QY 421 DIAYGAGRNVVYRAIKYDDYAKLFTFGSVADKGSATTFPVSGATFVTATLYMDTGSDDI 480
DB 421 DIAYGAGRNVVYRAIKYDDYAKLFTFGSVADKGSATTFPVSGATFVTATLYMDTGSDDI 480
QY 481 DLYLDPNGNEVDYSYTAAYGFEKVGYNPTAGTWTVKVSYKGAANYQVDVSDGSLSQ 540
DB 481 DLYLDPNGNEVDYSYTAAYGFEKVGYNPTAGTWTVKVSYKGAANYQVDVSDGSLSQ 540
QY 541 SGGGNPNPNPNPTPTDPTFTGTVNDYMDTFTFMANVSGATKITGDLTFDTSYND 600
DB 541 SGGGNPNPNPNPTPTDPTFTGTVNDYMDTFTFMANVSGATKITGDLTFDTSYND 600

QY 601 LDLYDPNGNLVDRSTSSNSYEHVEYANPAFGTWTFLVYASTYGMADYQLKAVVYVG 659
DB 601 LDLYDPNGNLVDRSTSSNSYEHVEYANPAFGTWTFLVYASTYGMADYQLKAVVYVG 659

RESULT 3

US-08-894-818B-35
; Sequence 35, Application US/088994818B
; Patent No. 6261822
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MITTA, Masanori
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,818B
; FILING DATE: 20-MAY-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03253
; FILING DATE: 07-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 323285/1995
; FILING DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: TAKAKURA=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 654 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-894-818B-35

Query Match 87.4%; Score 2997.5; DB 2; Length 654;
Best Local Similarity 88.3%; Pred. No. 8.5e-210;
Matches 582; Conservative 22; Mismatches 50; Indels 5; Gaps 1;

QY 1 MGKALKLILVILVLGLVGSVAAAPPEKKEQVRNVEKNYGLTPGLFRKIQKLNPEEIS 60
DB 1 MGKALKLILVILVLGLVGSVAAAPPEKKEQVRNVEKNYGLTPGLFRKIQKLNPEEIS 60
QY TVIVFENHREKEIAVRVLELMGAKRVYVYHIIIPAIADLKVRDLVLVLSGLTGKAKLSGV 120
DB TVIVFENHREKEIAVRVLELMGAKRVYVYHIIIPAIADLKVRDLVLVLSGLTGKAKLSGV 120
QY RFIOEDYKVTVALEGLDESAAQVMATYVWNLGYDGSGITIGIIDTGIDASHPDLOGKV 180
DB RFIOEDYKVTVALEGLDESAAQVMATYVWNLGYDGSGITIGIIDTGIDASHPDLOGKV 180

QY 181 ISWVDFVNGRSYPYDDHGHGHTVASIAAGTGAASNGKYKGMAPGAKIACIKVLGADGSGS 240
DB 181 ISWVDFVNGRSYPYDDHGHGHTVASIAAGTGAASNGKYKGMAPGAKIACIKVLGADGSGS 240
QY 241 ISTIIKGVMAVDNKKDKGIKVINSLGSSGSSDGTDSLQAVNNAMDGIYVCVAAGNS 300
DB 241 ISTIIKGVMAVDNKKDKGIKVINSLGSSGSSDGTDSLQAVNNAMDGIYVCVAAGNS 300
QY 301 GPNRTYVGSRAASKVITGAVDSDNNTASFSSRGPTADGRKPEVAVAPGVIIIPARASG 360
DB 301 GPNRTYVGSRAASKVITGAVDSDNNTASFSSRGPTADGRKPEVAVAPGVIIIPARASG 360
QY 361 TSMGTPINDYTTKASGTSMTAPHSVGVALIIQAPSPWPKVKTALLETADIVAPKEIA 420
DB 361 TSMGTPINDYTTKASGTSMTAPHSVGVALIIQAPSPWPKVKTALLETADIVAPKEIA 420
QY 421 DIAYGAGRVNVYKAIKYDDYAKLFTFGSVADKGSATHTFDVSGATFTVATLTLYMDTSSDI 480
DB 421 DIAYGAGRVNVYKAIKYDDYAKLFTFGSVADKGSATHTFDVSGATFTVATLTLYMDTSSDI 480
QY 481 DLVDPNGNEVDYSYTAAYGFEKGYVNPDTAGTWTVKVSYKGAANYVDVSDGSLG 540
DB 481 DLVDPNGNEVDYSYTAAYGFEKGYVNPDTAGTWTVKVSYKGAANYVDVSDGSLG 540
QY 541 SCGGNPNPNPNPPTTDTOTFTGTSVNDYMDTSDFTNANVSGATKINGDLTFDTSYND 600
DB 541 SCGGNPNPNPNPPTTDTOTFTGTSVNDYMDTSDFTNANVSGATKINGDLTFDTSYND 600
QY 541 -----PGSSPSPQPEFTVYDAKTFQSDHYHYSRDTFTMTVNSGATKITGDVFDTSYND 595
DB 541 -----PGSSPSPQPEFTVYDAKTFQSDHYHYSRDTFTMTVNSGATKITGDVFDTSYND 595
QY 601 LDLYDPNGNLVDRSTSSNSYEHVEYANPAFGTWTFLVYASTYGMADYQLKAVVYVG 659
DB 601 LDLYDPNGNLVDRSTSSNSYEHVEYANPAFGTWTFLVYASTYGMADYQLKAVVYVG 659

RESULT 4

US-09-445-472-16
; Sequence 16, Application US/09445472
; Patent No. 6358726
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/09/445,472
; CURRENT FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
; US-09-445-472-16

Query Match 87.4%; Score 2997.5; DB 2; Length 654;
Best Local Similarity 88.3%; Pred. No. 8.5e-210;
Matches 582; Conservative 22; Mismatches 50; Indels 5; Gaps 1;

QY 1 MGKALKLILVILVLGLVGSVAAAPPEKKEQVRNVEKNYGLTPGLFRKIQKLNPEEIS 60
DB 1 MGKALKLILVILVLGLVGSVAAAPPEKKEQVRNVEKNYGLTPGLFRKIQKLNPEEIS 60
QY TVIVFENHREKEIAVRVLELMGAKRVYVYHIIIPAIADLKVRDLVLVLSGLTGKAKLSGV 120
DB TVIVFENHREKEIAVRVLELMGAKRVYVYHIIIPAIADLKVRDLVLVLSGLTGKAKLSGV 120
QY RFIOEDYKVTVALEGLDESAAQVMATYVWNLGYDGSGITIGIIDTGIDASHPDLOGKV 180
DB RFIOEDYKVTVALEGLDESAAQVMATYVWNLGYDGSGITIGIIDTGIDASHPDLOGKV 180

QY 181 IGMVDFVNGRSYYPDDHGHGTHVASIAAGTGAASNGKYGMAFGAKLAGIKVLGADGSGS 240
DB 181 IGMVDFVNGRSYYPDDHGHGTHVASIAAGTGAASNGKYGMAFGAKLAGIKVLGADGSGS 240
QY 241 ISTIIKGVEMAVDNKCKYIGIKVINLSLGSQSSDGTDLGSAVNNAWDAGIYVCAAGNS 300
DB 241 ISTIIKGVEMAVDNKCKYIGIKVINLSLGSQSSDGTDLGSAVNNAWDAGIYVCAAGNS 300
QY 301 GPNNTYVSPAAASKVITVGAVDSDNIAASFSSRGPTADGRLKPEVVAPEGVDIAPRASG 360
DB 301 GPNNTYVSPAAASKVITVGAVDSDNIAASFSSRGPTADGRLKPEVVAPEGVDIAPRASG 360
QY 361 TSMGTPINDYTTKASGTSMATPHVSGVALLIOAHPSWTPDKVKTALLETADIYAPKEIA 420
DB 361 TSMGTPINDYTTKASGTSMATPHVSGVALLIOAHPSWTPDKVKTALLETADIYAPKEIA 420
QY 421 DIAYGAGRVNAYKAIKYDDYAKLTFTGSAVDKGSATHTPDVSGATFTVATLTYMDTGSDDI 480
DB 421 DIAYGAGRVNAYKAIKYDDYAKLTFTGSAVDKGSATHTPDVSGATFTVATLTYMDTGSDDI 480
QY 481 DLYYDPNGNEVDYSYTAAYGFEKVGYYNPTAGTWTVKVSYKGAANYQVDVSDGSLSQ 540
DB 481 DLYYDPNGNEVDYSYTAAYGFEKVGYYNPTAGTWTVKVSYKGAANYQVDVSDGSLSQ 540
QY 541 SGGGNPNPNPNPTPTDTQFTGSAVDYMDTSDFTMNVSAGTKITGDLTPDTSYND 600
DB 541 SGGGNPNPNPNPTPTDTQFTGSAVDYMDTSDFTMNVSAGTKITGDLTPDTSYND 600
QY 596 LDLYLDPNGNLDVDRSTSSNSYEHVEYANPAGTWTFLVYASTYGMADYOLKAVYVYG 659
DB 596 LDLYLDPNGNLDVDRSTSSNSYEHVEYANPAGTWTFLVYASTYGMADYOLKAVYVYG 654

RESULT 5

US-10-090-624-16
Sequence 16, Application US/10090624
Patent No. 6783970
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOJO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA=6
CURRENT APPLICATION NUMBER: US/10/090,624
PRIOR FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: 09/445,472
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patentin version 3.0
SEQ ID NO 16
LENGTH: 654
TYPE: PRT
ORGANISM: Pyrococcus furiosus
US-10-090-624-16

Query Match 87.4%; Score 2997.5; DB 2; Length 654;
Best Local Similarity 88.3%; Prid. No. 8.5e-10;
Matches 582; Conservative 22; Mismatches 50; Indels 5; Gaps 1;

QY 1 MKGALKALIVITVIGLVGSAVAAPKRYQVAVNEKNGGLTPGLFRKIQKLNPEEIS 60
DB 1 MKGALKALIVITVIGLVGSAVAAPKRYQVAVNEKNGGLTPGLFRKIQKLNPEEIS 60
QY 61 TVIVFENHREKEIAVAVLEIMGAKVYVYHIIIPAIADLKVDLIVISGLTGKAKLSGV 120
DB 61 TVIVFENHREKEIAVAVLEIMGAKVYVYHIIIPAIADLKVDLIVISGLTGKAKLSGV 120
QY 121 RFIOEDYKTVSALEBESDAQAQVMAITYVMNLGYGSGITIGIIDTGDASHPDLQGRV 180

DB 121 RFIOEDYKTVSALEBESDAQAQVMAITYVMNLGYGSGITIGIIDTGDASHPDLQGRV 180
QY 181 IGMVDFVNGRSYYPDDHGHGTHVASIAAGTGAASNGKYGMAFGAKLAGIKVLGADGSGS 240
DB 181 IGMVDFVNGRSYYPDDHGHGTHVASIAAGTGAASNGKYGMAFGAKLAGIKVLGADGSGS 240
QY 241 ISTIIKGVEMAVDNKCKYIGIKVINLSLGSQSSDGTDLGSAVNNAWDAGIYVCAAGNS 300
DB 241 ISTIIKGVEMAVDNKCKYIGIKVINLSLGSQSSDGTDLGSAVNNAWDAGIYVCAAGNS 300
QY 301 GPNNTYVSPAAASKVITVGAVDSDNIAASFSSRGPTADGRLKPEVVAPEGVDIAPRASG 360
DB 301 GPNNTYVSPAAASKVITVGAVDSDNIAASFSSRGPTADGRLKPEVVAPEGVDIAPRASG 360
QY 361 TSMGTPINDYTTKASGTSMATPHVSGVALLIOAHPSWTPDKVKTALLETADIYAPKEIA 420
DB 361 TSMGTPINDYTTKASGTSMATPHVSGVALLIOAHPSWTPDKVKTALLETADIYAPKEIA 420
QY 421 DIAYGAGRVNAYKAIKYDDYAKLTFTGSAVDKGSATHTPDVSGATFTVATLTYMDTGSDDI 480
DB 421 DIAYGAGRVNAYKAIKYDDYAKLTFTGSAVDKGSATHTPDVSGATFTVATLTYMDTGSDDI 480
QY 481 DLYYDPNGNEVDYSYTAAYGFEKVGYYNPTAGTWTVKVSYKGAANYQVDVSDGSLSQ 540
DB 481 DLYYDPNGNEVDYSYTAAYGFEKVGYYNPTAGTWTVKVSYKGAANYQVDVSDGSLSQ 540
QY 541 SGGGNPNPNPNPTPTDTQFTGSAVDYMDTSDFTMNVSAGTKITGDLTPDTSYND 600
DB 541 SGGGNPNPNPNPTPTDTQFTGSAVDYMDTSDFTMNVSAGTKITGDLTPDTSYND 600
QY 596 LDLYLDPNGNLDVDRSTSSNSYEHVEYANPAGTWTFLVYASTYGMADYOLKAVYVYG 659
DB 596 LDLYLDPNGNLDVDRSTSSNSYEHVEYANPAGTWTFLVYASTYGMADYOLKAVYVYG 654

RESULT 6

US-09-841-553-35
Sequence 35, Application US/09841553
Patent No. 6849441
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MITTA, Masamori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSER: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/841,553
FILING DATE: 24-Apr-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/894,818
FILING DATE: <Unknown>
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618

REFERENCE/DOCKET NUMBER: TAKAKURA=1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-5197
 TELEFAX: (202) 737-3528
 INFORMATION FOR SEQ ID NO: 35:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 654 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 35:
 US-09-841-553-35

Query Match 87.4%; Score 2997.5; DB 2; Length 654;
 Best Local Similarity 88.3%; Pred. No. 8.5e-210;
 Matches 582; Conservative 22; Mismatches 50; Indels 5; Gaps 1;

QY 1 MKGLKALILVILVLGLVGSVAAPAEKVEQVNNVEKNYGLTPGLFRKIQKLNPEEIS 60
 DB 1 MKGLKALILVILVLGLVGSVAAPAEKVEQVNNVEKNYGLTPGLFRKIQKLNPEEIS 60
 QY 61 TVIVFENHREKEIAVAVLELMGAKVRYVYHIIPALADLKVDLVISGLTSGKATLSCV 120
 DB 61 TVIVFENHREKEIAVAVLELMGAKVRYVYHIIPALADLKVDLVISGLTSGKATLSCV 120
 QY 121 RPIQDEYKTVASLEGLDESAQVMAATVYMNIGYSGSGITIGIDTGDASHPDLOGKV 180
 DB 121 RPIQDEYKTVASLEGLDESAQVMAATVYMNIGYSGSGITIGIDTGDASHPDLOGKV 180
 QY 181 IGVAVDFVNGRSYYPYDHGHTHVASIAAGTGAASNGKYGMAPGAKIAGIKVLGADGSGS 240
 DB 181 IGVAVDFVNGRSYYPYDHGHTHVASIAAGTGAASNGKYGMAPGAKIAGIKVLGADGSGS 240
 QY 241 ISTIIGVEWAVNDKDKYGIKINLSLGSQSSDGTDSLQAVNNAMDAGIIVCVAAQNS 300
 DB 241 ISTIIGVEWAVNDKDKYGIKINLSLGSQSSDGTDSLQAVNNAMDAGIIVCVAAQNS 300
 QY 301 GPRYTVVGSPPAAASKYITVCAVDSNDNIASFSSRGPTADGRLKPEVVAQGVDIIAARASG 360
 DB 301 GPRYTVVGSPPAAASKYITVCAVDSNDNIASFSSRGPTADGRLKPEVVAQGVDIIAARASG 360
 QY 361 TSMGQPIINDYTTASGSMATPHVSGVGLILQAHPSWTPDKYKTALEIETADIVAPKEIA 420
 DB 361 TSMGQPIINDYTTASGSMATPHVSGVGLILQAHPSWTPDKYKTALEIETADIVAPKEIA 420
 QY 421 DIAVAGRVNVAIKYDYAKLTFTGSAVDKGSATHTPDVSGATEVTATLYWDTGSSDI 480
 DB 421 DIAVAGRVNVAIKYDYAKLTFTGSAVDKGSATHTPDVSGATEVTATLYWDTGSSDI 480
 QY 481 DLVLYDPNGNEVDYSTAYYGFEEKVGYNFTAGTWTVKVSYKGAANYQVDVSDGSLQ 540
 DB 481 DLVLYDPNGNEVDYSTAYYGFEEKVGYNFTAGTWTVKVSYKGAANYQVDVSDGSLQ 540
 QY 541 SGGGNNPNPNPPTPTTQTFGTGSVNDWDSDPTTMVNSGAKRTIGDLEFDTSYND 600
 DB 541 SGGGNNPNPNPPTPTTQTFGTGSVNDWDSDPTTMVNSGAKRTIGDLEFDTSYND 600
 QY 601 LDLYLYDPNGNLVDRSTSSNSYEHVEYANPAPGTMTFLVYAYSTYGADYQLKAVVYGG 659
 DB 596 LDLYLYDPNGNLVDRSESPNSYEHVEYLTLPAPGTWFLVYAYTYTGKAYIELTAKYYTG 654

RESULT 7
 US-08-894-818B-1
 ; Sequence 1, Application US/0894818B
 ; Patent No. 6261822
 ; GENERAL INFORMATION:
 ; APPLICANT: TAKAKURA, Hikaru
 ; APPLICANT: MORISHITA, Mio
 ; APPLICANT: YAMAMOTO, Katsuhiko
 ; APPLICANT: MITTA, Masanori
 ; APPLICANT: ASADA, Kiyozo

APPLICANT: TSUNASAWA, Susumu
 APPLICANT: KATO, Ikunoshin
 TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Browdy and Neimark
 STREET: 419 Seventh Street N.W., Ste. 300
 CITY: Washington
 STATE: D.C.
 COUNTRY: United States of America
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA: US/08/894,818B
 APPLICATION NUMBER: 25,618
 FILING DATE: 12-DEC-1995
 CLASSIFICATION: 435
 PRIORITY DATE: 07-NOV-1996
 PCT/JP96/03253
 APPLICATION NUMBER: 323285/1995
 FILING DATE: 12-DEC-1995
 NAME: Browdy, Roger L.
 REGISTRATION NUMBER: 25,618
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-5197
 TELEFAX: (202) 737-3528
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 659 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-894-818B-1

Query Match 85.0%; Score 2914; DB 2; Length 659;
 Best Local Similarity 83.7%; Pred. No. 1e-203;
 Matches 558; Conservative 47; Mismatches 46; Indels 16; Gaps 7;

QY 1 MKGLKALILVILVLGLVGSVAAPAEKVEQVNN--VEKNYGLTPGLFRKIQKLNPE 57
 DB 1 MKGLKALILVILVLGLVGLAGTALAAPVKV--VRNNAVQOKNYGLTPGLFRKIQKLNPE 58
 QY 58 EISTVIVFENHREKEIAVAVLELMGAKVRYVYHIIPALADLKVDLVISGL--TG--G 113
 DB 59 EVDVTFMFSSYDRBRAVAVLELMGAKVRYVYHIIPALADLKVDLVISGL--TG--G 118
 QY 114 KAKLSGVRIOEDYKTVASLEGLDESAQVMAATVYMNIGYSGSGITIGIDTGDASH 172
 DB 119 NTRVSGIKTIOEDYKTVASLEGLDESAQVMAATVYMNIGYSGSGITIGIDTGDASH 173
 QY 173 HPDLQKVGIVGWDFVNGRSYYPYDHGHTHVASIAAGTGAASNGKYGMAPGAKIAGIRV 232
 DB 174 HEDLKGVIGWDAVNGRSTPYDDGHTHVGIVAGTGSV--NSQVIGVAPGKLVGVKV 232
 QY 233 LGADSGSISTIIKGYEAVNDKDKYGIKINLSLGSQSSDGTDSLQAVNNAMDAGIV 292
 DB 233 LGADSGSVSTIIAGDWWQNKDKYGIKINLSLGSQSSDGTDSLQAVNNAMDAGIV 292
 QY 293 VCVAAAGSPNTYTVGSPAAASKYITVCAVDSNDNIASFSSRGPTADGRLKPEVVAQGV 352
 DB 293 VCVAAAGSPNTYTVGSPAAASKYITVCAVDSNDNIASFSSRGPTADGRLKPEVVAQGV 352
 QY 353 IIAPRASGTSMTGPIINDYTTKASGSMATPHVSGVGLILQAHPSWTPDKYKTALEIETAD 412
 DB 353 IIAPRASGTSMTGPIINDYTTKASGSMATPHVSGVGLILQAHPSWTPDKYKTALEIETAD 412


```
QY 413 IYAPKEIADIAYAGAGRVNYKAIKYDDYAKLFTFGSVADKGSATHTFDVSGATFTATLY 472
Db 413 IYAPKEIADIAYAGAGRVNYKAIKYDDYAKLFTFGSVADKGSATHTFDVSGATFTATLY 472
QY 473 WDTGSSDIDLXYLDPNGENVDYSYTAAYGFEKVGYNPTAGTWTVAKVSYKGAANYQYDV 532
Db 473 WDTGSSDIDLXYLDPNGENVDYSYTAAYGFEKVGYNPTAGTWTVAKVSYKGAANYQYDV 532
QY 533 VSDGSLSGSGGPNPNPNPNPTPTTDTQTFGTVNDWDTSDPTTMVNSGATKTYGDL 592
Db 533 VSDGSLSGSGGPNPNPNPNPTPTTDTQTFGTVNDWDTSDPTTMVNSGATKTYGDL 592
QY 593 TFDTSYNDLXYLDPNGLVDRTSSNSYEHVEYANPAPGTMTFLVAYSTYGMADYQL 652
Db 593 TFDTSYNDLXYLDPNGLVDRTSSNSYEHVEYANPAPGTMTFLVAYSTYGMADYQL 652
QY 653 KAVVYVG 659
Db 653 KAVVYVG 659
```

RESULT 8

```
US-09-445-472-12
; Sequence 12, Application US/09445472
; Patent No. 6358726
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunobu
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/09/445.472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Thermococcus celer
US-09-445-472-12
```

Query Match 85.0%; Score 2914; DB 2; Length 659;

Best Local Similarity 83.7%; Pred. No. 1e-203;

Matches 558; Conservative 47; Mismatches 46; Indels 16; Gaps 7;

```
QY 1 MKGLKALILVILVGLVYGVAAAPKKVEQVRN---VEKNYGLTFPGLFKRIQKLNPN 57
Db 1 MKRLGAVVLAIVVGLAGTALAAPVFPV--VRNNAVQCKNYGLTFGLFKKVRNMNQ 58
QY 58 EISTVIYFENHREKEIAVRVLELMGAKRVYVHIIPALADLKVRDLVLSGL--TG--G 113
Db 58 EIVDTIYFSGYGRDRAVKVRLMGAQVKSYKIIIPAVAKKIARDLLLAGMIDTGYFG 118
QY 114 KAKLSGVRFIOEDYKVTYSAELGLDESAQVMAITYWN--LGYDGSGITIGIIDTGIDAS 172
Db 119 NTRVSGIFIOEDYKQVDDA-----TSVQIGADTYMNSLGYDGSVVVAIVDTGIDAN 173
QY 173 HPDLQGVIGVWDFVNGRSYYPYDDHGHGHTVASTAAGTGAASNGKYKGMAFGAKLAGIV 232
Db 174 HPDLKGVIGVWDAVNGRSTPYDDQGHGHTVAGIVAGTGSV--NSQYIGVAPGAKLVGVKV 232
QY 233 LGADGSGSISTIIKGYEMAVDNKDKYKIVINLSLSSGSSDGTDSLGAANNAMAGIV 292
Db 233 LGADGSGSVSTIIAGVDMVYQNDKYGIRVINLSLSSGSSDGTDSLGAANNAMAGIV 292
QY 293 VCVAAAGNSGNTYVSSPAASRVITVGAVDSDNNTIASFSSRGPTADGRLKPEVAPGYD 352
Db 293 VCVAAAGNSGNTYVSSPAASRVITVGAVDSDNNTIASFSSRGPTADGRLKPEVAPGYD 352
```

```
QY 353 IIAPRASGTSMTPIINDYTTKASGTSMAHPVSGVALLIIQAHPSWTPDKVTALLIETAD 412
Db 353 IIAPRASGTSMTPIINDYTTKASGTSMAHPVSGVALLIIQAHPSWTPDKVTALLIETAD 412
QY 413 IYAPKEIADIAYAGAGRVNYKAIKYDDYAKLFTFGSVADKGSATHTFDVSGATFTATLY 472
Db 413 IYAPKEIADIAYAGAGRVNYKAIKYDDYAKLFTFGSVADKGSATHTFDVSGATFTATLY 472
QY 473 WDTGSSDIDLXYLDPNGENVDYSYTAAYGFEKVGYNPTAGTWTVAKVSYKGAANYQYDV 532
Db 473 WDTGSSDIDLXYLDPNGENVDYSYTAAYGFEKVGYNPTAGTWTVAKVSYKGAANYQYDV 532
QY 533 VSDGSLSGSGGPNPNPNPNPTPTTDTQTFGTVNDWDTSDPTTMVNSGATKTYGDL 592
Db 533 VSDGSLSGSGGPNPNPNPNPTPTTDTQTFGTVNDWDTSDPTTMVNSGATKTYGDL 592
QY 593 TFDTSYNDLXYLDPNGLVDRTSSNSYEHVEYANPAPGTMTFLVAYSTYGMADYQL 652
Db 593 TFDTSYNDLXYLDPNGLVDRTSSNSYEHVEYANPAPGTMTFLVAYSTYGMADYQL 652
QY 653 KAVVYVG 659
Db 653 KAVVYVG 659
```

RESULT 9

```
US-10-090-624-12
; Sequence 12, Application US/10090624
; Patent No. 6783970
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunobu
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/10/090.624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445.472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Thermococcus celer
US-10-090-624-12
```

Query Match 85.0%; Score 2914; DB 2; Length 659;

Best Local Similarity 83.7%; Pred. No. 1e-203;

Matches 558; Conservative 47; Mismatches 46; Indels 16; Gaps 7;

```
QY 1 MKGLKALILVILVGLVYGVAAAPKKVEQVRN---VEKNYGLTFPGLFKRIQKLNPN 57
Db 1 MKRLGAVVLAIVVGLAGTALAAPVFPV--VRNNAVQCKNYGLTFGLFKKVRNMNQ 58
QY 58 EISTVIYFENHREKEIAVRVLELMGAKRVYVHIIPALADLKVRDLVLSGL--TG--G 113
Db 58 EIVDTIYFSGYGRDRAVKVRLMGAQVKSYKIIIPAVAKKIARDLLLAGMIDTGYFG 118
QY 114 KAKLSGVRFIOEDYKVTYSAELGLDESAQVMAITYWN--LGYDGSGITIGIIDTGIDAS 172
Db 119 NTRVSGIFIOEDYKQVDDA-----TSVQIGADTYMNSLGYDGSVVVAIVDTGIDAN 173
QY 173 HPDLQGVIGVWDFVNGRSYYPYDDHGHGHTVASTAAGTGAASNGKYKGMAFGAKLAGIV 232
Db 174 HPDLKGVIGVWDAVNGRSTPYDDQGHGHTVAGIVAGTGSV--NSQYIGVAPGAKLVGVKV 232
QY 233 LGADGSGSISTIIKGYEMAVDNKDKYKIVINLSLSSGSSDGTDSLGAANNAMAGIV 292
```


APPLICANT: TSUNASAWA, Sueumu
APPLICANT: KATO, Ikumohin
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,818B
FILING DATE: 20-MAY-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03253
FILING DATE: 07-NOV-1996
PRIOR APPLICATION DATA: JP 323285/1995
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 522 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: /note= Xaa at position 428 is Gly or Val.
US-08-894-818B-3

Query Match 68.5%; Score 2349.5; DB 2; Length 522;
Best Local Similarity 85.2%; Pred. No. 9.1e-163;
Matches 449; Conservative 22; Mismatches 51; Indels 5; Gaps 1;

QY 133 AELEGLDESAQAQMAYVWNLGYDGSGITIGIIDTGIDASHPDLQKVIQWVDFVNGRSY 192
DB 1 AELEGLDESAQAQMAYVWNLGYDGSGITIGIIDTGIDASHPDLQKVIQWVDFVNGRSY 60
QY 193 PYDDHGHGTHVASIAAGTGAASNGKYGMAFGAKLAGIKVLGADSGSISTIIKGVEMAV 252
DB 61 PYDDHGHGTHVASIAAGTGAASNGKYGMAFGAKLAGIKVLGADSGSISTIIKGVEMAV 120
QY 253 DNKDKYKIVINLSIGSSQSSDGTSLGQAVNNAVDAGIVVCAAGNSGPNITVYVSPA 312
DB 121 DNKDKYKIVINLSIGSSQSSDGTSLGQAVNNAVDAGIVVCAAGNSGPNITVYVSPA 180
QY 313 ASKVIITVGAVDNDINIASFSRGPPTADGRKPEVVAFGVDIIAPASGTSMTPIINDYTT 372
DB 181 ASKVIITVGAVDKDYVITFSRSGPTADGRKPEVVAFGVWIIAARASGTSMTQPIINDYTT 240
QY 373 KASGTSMTPIHVSQVAGALLLOAHPSWTPDKVKTALIFETADIYAPKEIADIAGAGRVNY 432
DB 241 AAPGTSMTPIHVSQVAGALLLOAHPSWTPDKVKTALIFETADIYAPKEIADIAGAGRVNY 300
QY 433 KAIKVDYAKLFTGQSVADKSGATHTFPVSGATFTVATLTYMDTSSDIDLVIYDPNGNEV 492
DB 301 KAINYDNYAKLVFTGYVANKSGQTHQFVLSGASFTATLYMDNANSDLDLVIYDPNGNQV 360
QY 493 DYSYTAIYGFEEKYGYNNPTAGTWTAVVSVYKGAANYQVDVVSQSGSGGNPNPNPNP 552

DB 361 DYSYTAIYGFEEKYGYNNPTAGTWTIKVSGSAGANYQVDVVSQSGSLQ-----PGSSPSF 415
QY 553 NPTPTDTQTFGQSVVDYMDTSDPTFMNNSGATKLTGDLTPTSTNDLVIYDPNGNL 612
DB 416 QPEPTVDAKTFQXSDHYHYDRSDPTFMNNSGATKLTGDLVPTSTYHDLDLVIYDPNGNL 475
QY 613 VDRSTSSNSYEHVEYANPAPGTWTFVVAYSTYGMADYOLKAVVYVY 659
DB 476 VDRSESPNSYEHVEYLTAPAGTWTFVVAIYTYGMAVEELTKVYVY 522

RESULT 12
US-09-445-472-4
Sequence 4, Application US/09445472
Patent No. 6358726
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOJO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikumohin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA=6
CURRENT APPLICATION NUMBER: US/09/445,472
CURRENT FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patent In version 3.0
SEQ ID NO 4
LENGTH: 522
TYPE: PRT
ORGANISM: Pyrococcus furiosus
FEATURE:
NAME/KEY: misc feature
LOCATION: (428)..(428)
OTHER INFORMATION: Xaa at position 428 is Gly or Val.
US-09-445-472-4

Query Match 68.5%; Score 2349.5; DB 2; Length 522;
Best Local Similarity 85.2%; Pred. No. 9.1e-163;
Matches 449; Conservative 22; Mismatches 51; Indels 5; Gaps 1;

QY 133 AELEGLDESAQAQMAYVWNLGYDGSGITIGIIDTGIDASHPDLQKVIQWVDFVNGRSY 192
DB 1 AELEGLDESAQAQMAYVWNLGYDGSGITIGIIDTGIDASHPDLQKVIQWVDFVNGRSY 60
QY 193 PYDDHGHGTHVASIAAGTGAASNGKYGMAFGAKLAGIKVLGADSGSISTIIKGVEMAV 252
DB 61 PYDDHGHGTHVASIAAGTGAASNGKYGMAFGAKLAGIKVLGADSGSISTIIKGVEMAV 120
QY 253 DNKDKYKIVINLSIGSSQSSDGTSLGQAVNNAVDAGIVVCAAGNSGPNITVYVSPA 312
DB 121 DNKDKYKIVINLSIGSSQSSDGTSLGQAVNNAVDAGIVVCAAGNSGPNITVYVSPA 180
QY 313 ASKVIITVGAVDNDINIASFSRGPPTADGRKPEVVAFGVDIIAPASGTSMTPIINDYTT 372
DB 181 ASKVIITVGAVDKDYVITFSRSGPTADGRKPEVVAFGVWIIAARASGTSMTQPIINDYTT 240
QY 373 KASGTSMTPIHVSQVAGALLLOAHPSWTPDKVKTALIFETADIYAPKEIADIAGAGRVNY 432
DB 241 AAPGTSMTPIHVSQVAGALLLOAHPSWTPDKVKTALIFETADIYAPKEIADIAGAGRVNY 300
QY 433 KAIKVDYAKLFTGQSVADKSGATHTFPVSGATFTVATLTYMDTSSDIDLVIYDPNGNEV 492
DB 301 KAINYDNYAKLVFTGYVANKSGQTHQFVLSGASFTATLYMDNANSDLDLVIYDPNGNQV 360
QY 493 DYSYTAIYGFEEKYGYNNPTAGTWTAVVSVYKGAANYQVDVVSQSGSGGNPNPNPNP 552
DB 361 DYSYTAIYGFEEKYGYNNPTDGTWITIKVSYSGSANYQVDVVSQSGSLQ-----PGSSPSF 415
QY 553 NPTPTDTQTFGQSVVDYMDTSDPTFMNNSGATKLTGDLTPTSTNDLVIYDPNGNL 612

Db 416 QPEPTADAKTFQXSDHYDSDTFTMTVNSGATKITGDLVFDTSVHDDLVLVDPNOKL 475
QY 613 VDRSTSSNGEYEHYANPAGTMTFLVYASTYGADYOLKAVVYVG 659
Db 476 VDRSESPNSYEYEHVILTPAGTMYFLVYAYTYGMAVYELTAKVYVG 522

RESULT 13

US-10-090-624-4
Sequence 4, Application US/10090624
Patent No. 6783970
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOTO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA=6
CURRENT APPLICATION NUMBER: US/10/090,624
PRIOR FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: 09/445,472
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patentin version 3.0
SEQ ID NO 4
LENGTH: 522
TYPE: PRT
ORGANISM: Pyrococcus furiosus
FEATURE:
NAME/KEY: misc.feature
LOCATION: (428)..(428)
OTHER INFORMATION: Xaa at position 428 is Gly or Val.
US-10-090-624-4

Query Match 68.5%; Score 2349.5; DB 2; Length 522;
Best Local Similarity 85.2%; Pred. No. 9.1e-163;
Matches 449; Conservative 22; Mismatches 51; Indels 5; Gaps 1;

QY 133 AELEGDESAOVMATYVNNLGYDGSGITIGITIGIDGASHPDLQGVIGWDFVNGRSY 192
Db 1 AELEGDESAOVMATYVNNLGYDGSGITIGITIGIDGASHPDLQGVIGWDFVNGRSY 60
QY 193 PYDDHGHGTHVASTAGTGAASNGKTKGMAFGAKLAGIKVILGADSGSISTIIKGVEMAV 252
Db 61 PYDDHGHGTHVASTAGTGAASNGKTKGMAFGAKLAGIKVILGADSGSISTIIKGVEMAV 120
QY 253 DNKDKYGIKVINLSLSSQSSDGTDSLGAQVNNAMDAIVCVAAAGNSGPNITYVGSPPA 312
Db 121 DNKDKYGIKVINLSLSSQSSDGTDSLGAQVNNAMDAIVCVAAAGNSGPNITYVGSPPA 180
QY 313 ASKVITVGAVDSDNITASPSRGPTADGRKPEPVNAPGVDTIAPRASGSMGTPIINDYTT 372
Db 181 ASKVITVGAVDKDYITTSRSRPTADGRKPEPVNAPGVDTIAPRASGSMGTPIINDYTT 240
QY 373 KASGTSMATPHVSGVALILQAPSWTPDKVKTALLETADIIVAPKEIADIAYAGRVNYY 432
Db 241 AARGTSMATPHVAGIALILQAPSWTPDKVKTALLETADIIVAPKEIADIAYAGRVNYY 300
QY 433 KAIKYDDYAKLFTTSGVADKGSATHTFDVSGATFTATTLTMDTSSDIDLVLVDPNANEV 492
Db 301 KAINYNNYAKLFTTSGVADKGSATHTFDVSGATFTATTLTMDTSSDIDLVLVDPNANEV 360
QY 493 DVSATYAYGPEKGYNPTAGTWTIVVSYKGAANOVDVSDSLSLSSGSGGPNPNPNPN 552
Db 361 DVSATYAYGPEKGYNPTAGTWTIVVSYKGAANOVDVSDSLSLSSGSGGPNPNPNPN 415
QY 553 NPTPTDQTFQTSVNDYWDYDSTFTMTVNSGATKITGDLTFTDTSYNDLVLVDPNGNL 612
Db 416 QPEPTADAKTFQXSDHYDSDTFTMTVNSGATKITGDLVFDTSVHDDLVLVDPNOKL 475

QY 613 VDRSTSSNGEYEHYANPAGTMTFLVYASTYGADYOLKAVVYVG 659
Db 476 VDRSESPNSYEYEHVILTPAGTMYFLVYAYTYGMAVYELTAKVYVG 522

RESULT 14

US-09-841-553-3
Sequence 3, Application US/09841553
Patent No. 6849441
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MURTA, Masanori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
KATO, Ikunoshin
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/841,553
FILING DATE: 24-Apr-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/894,818
FILING DATE: <Unknown>
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 522 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
OTHER INFORMATION:
Xaa at position 428 is Gly or Val.
US-09-841-553-3

Query Match 68.5%; Score 2349.5; DB 2; Length 522;
Best Local Similarity 85.2%; Pred. No. 9.1e-163;
Matches 449; Conservative 22; Mismatches 51; Indels 5; Gaps 1;

QY 133 AELEGDESAOVMATYVNNLGYDGSGITIGITIGIDGASHPDLQGVIGWDFVNGRSY 192
Db 1 AELEGDESAOVMATYVNNLGYDGSGITIGITIGIDGASHPDLQGVIGWDFVNGRSY 60
QY 193 PYDDHGHGTHVASTAGTGAASNGKTKGMAFGAKLAGIKVILGADSGSISTIIKGVEMAV 252
Db 61 PYDDHGHGTHVASTAGTGAASNGKTKGMAFGAKLAGIKVILGADSGSISTIIKGVEMAV 120
QY 253 DNKDKYGIKVINLSLSSQSSDGTDSLGAQVNNAMDAIVCVAAAGNSGPNITYVGSPPA 312

```

Db 121 DNKRYGKIVNLISGSSQSDGTDALSOAVNAAMDAGLVVVAAGNSGPKYTIIGSPA 180
QY 313 ASKVITVGAVDSDNINISFSSRGPADGRLEKEVVAAPVDIIAARAGTSMGTINNYT 372
Db 181 ASKVITVGAVDKIVITISFSSRGPADGRLEKEVVAAPVDIIAARAGTSMGTINNYT 240
QY 373 KASGTSMATPHVSGVAGLIIQAHPSMTPEKVKALIIETADIAPKEIADIAGAGRVNVY 432
Db 241 AAPGTSMATPHVAGIAALLQAHPSMTPEKVKALIIETADIAPKEIADIAGAGRVNVY 300
QY 433 KAIKIDYAKLITFTGSVADKGSATHTFDVSGATVATLTLYWDGSSDILYLDPNNGEV 492
Db 301 KAINVDNYAKLVFTGYVANKGSQTHOFVISCASFVATLTLYWDNANSDDLILYLDPNNGV 360
QY 493 DYSTAYYGFKEKVGYNPTAGTWTVKVVSYGKANYOVVSDGSLSSQSGGPNPNP 552
Db 361 DYSTAYYGFKEKVGYNPTAGTWTIKVVSISGSANYOVVSDGSLSSQSGGPNPNP 415
QY 553 NPTFTDTQFTGSDVNDYDTSPTFTMNVSAGATKIGDLTFTDTSYNDLILYLDPNNG 612
Db 416 QPEPTVAKTFOKSDHYRSDPTFTMTVNSGATKIGDLVFTDTSYNDLILYLDPNNG 475
QY 613 VDRSTSGSYEHVYANPAPGTWELVYASTYGMADYOLKAVVY 659
Db 476 VDRSTSGSYEHVYALTPAGTWTFLVAYTYGMAYELTAKVY 522

```

```

RESULT 15
US-09-445-472-1
; Sequence 1, Application US/09445472
; Patent No. 6358726
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunobu
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/09/445,472
; CURRENT FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 412
; TYPE: PRF
; ORGANISM: Pyrococcus furiosus
US-09-445-472-1

```

```

Query Match 56.0%; Score 1918; DB 2; Length 412;
Best Local Similarity 89.5%; Pred. No. 1.7e-131;
Matches 367; Conservative 15; Mismatches 28; Indels 0; Gaps 0;

```

```

QY 133 ALEEGIDESAQAQMATYVWNLGYDGGITIGIITDGDASHPLDQKVTIGWDFVNGRSY 192
Db 1 ALEEGIDESAQAQMATYVWNLGYDGGITIGIITDGDASHPLDQKVTIGWDFVNGRSY 60
QY 193 PYDDHGHGTHVASIAAGTGAASNGKYGAPAGKLAGIKVILGADSGSSTIIKGVEMAV 252
Db 61 PYDDHGHGTHVASIAAGTGAASNGKYGAPAGKLAGIKVILGADSGSSTIIKGVEMAV 120
QY 253 DNKRYGKIVNLISGSSQSDGTDLSQAVNNAAMDAGIVVCAAGNSGPNNTYVGSPPA 312
Db 121 DNKRYGKIVNLISGSSQSDGTDLSQAVNNAAMDAGIVVCAAGNSGPNNTYVGSPPA 180
QY 313 ASKVITVGAVDSDNINISFSSRGPADGRLEKEVVAAPVDIIAARAGTSMGTINNYT 372
Db 181 ASKVITVGAVDKIVITISFSSRGPADGRLEKEVVAAPVDIIAARAGTSMGTINNYT 240
QY 373 KASGTSMATPHVSGVAGLIIQAHPSMTPEKVKALIIETADIAPKEIADIAGAGRVNVY 432

```

```

Db 241 AAPGTSMATPHVAGIAALLQAHPSMTPEKVKALIIETADIAPKEIADIAGAGRVNVY 300
QY 433 KAIKIDYAKLITFTGSVADKGSATHTFDVSGATVATLTLYWDGSSDILYLDPNNGEV 492
Db 301 KAINVDNYAKLVFTGYVANKGSQTHOFVISCASFVATLTLYWDNANSDDLILYLDPNNGV 360
QY 493 DYSTAYYGFKEKVGYNPTAGTWTVKVVSYGKANYOVVSDGSLSSQSG 542
Db 361 DYSTAYYGFKEKVGYNPTAGTWTIKVVSISGSANYOVVSDGSLSSQSG 410

```

Search completed: January 6, 2007, 22:33:38
Job time : 38 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2007 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2007, 22:32:37 ; Search time 123.5 Seconds
(without alignments)
2471.729 Million cell updates/sec

Title: US-10-800-684-1
Perfect score: 3437
Sequence: 1 MKRLGAVLAVLVGLAGT.....YASTYGMADYQLKAVVYG 659

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA_Main:
1: /EMC_Celerra_SIDS3/ptocdata/2/pubpaa/US07_PUBCOMB.pep:*
2: /EMC_Celerra_SIDS3/ptocdata/2/pubpaa/US08_PUBCOMB.pep:*
3: /EMC_Celerra_SIDS3/ptocdata/2/pubpaa/US09_PUBCOMB.pep:*
4: /EMC_Celerra_SIDS3/ptocdata/2/pubpaa/US10A_PUBCOMB.pep:*
5: /EMC_Celerra_SIDS3/ptocdata/2/pubpaa/US10B_PUBCOMB.pep:*
6: /EMC_Celerra_SIDS3/ptocdata/2/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|---------------------|-------------------|
| 1 | 3437 | 100.0 | 659 | US-09-841-553-1 | Sequence 1, Appl |
| 2 | 3437 | 100.0 | 659 | US-10-090-624-12 | Sequence 12, Appl |
| 3 | 3437 | 100.0 | 659 | US-10-800-684-1 | Sequence 1, Appl |
| 4 | 3437 | 100.0 | 659 | US-10-888-588-12 | Sequence 12, Appl |
| 5 | 2914 | 84.8 | 659 | US-09-841-553-5 | Sequence 5, Appl |
| 6 | 2914 | 84.8 | 659 | US-10-800-684-5 | Sequence 5, Appl |
| 7 | 2483.5 | 72.3 | 654 | US-09-841-553-35 | Sequence 35, Appl |
| 8 | 2483.5 | 72.3 | 654 | US-10-090-624-16 | Sequence 16, Appl |
| 9 | 2483.5 | 72.3 | 654 | US-10-800-684-35 | Sequence 35, Appl |
| 10 | 2483.5 | 72.3 | 654 | US-10-888-588-16 | Sequence 16, Appl |
| 11 | 2138.5 | 62.2 | 522 | US-09-841-553-3 | Sequence 3, Appl |
| 12 | 2138.5 | 62.2 | 522 | US-10-090-624-4 | Sequence 4, Appl |
| 13 | 2138.5 | 62.2 | 522 | US-10-800-684-3 | Sequence 3, Appl |
| 14 | 2138.5 | 62.2 | 522 | US-10-888-588-4 | Sequence 4, Appl |
| 15 | 1707 | 49.7 | 412 | US-10-090-624-1 | Sequence 1, Appl |
| 16 | 1707 | 49.7 | 412 | US-10-888-588-1 | Sequence 1, Appl |
| 17 | 668.5 | 19.5 | 1139 | US-10-156-761-10656 | Sequence 10656, A |
| 18 | 657.5 | 19.1 | 1208 | US-10-156-761-13251 | Sequence 13251, A |
| 19 | 636 | 18.5 | 1079 | US-10-112-488-39 | Sequence 39, Appl |
| 20 | 636 | 18.5 | 1079 | US-10-673-860-8 | Sequence 8, Appl |
| 21 | 632.5 | 18.4 | 1237 | US-10-314-657-4 | Sequence 4, Appl |
| 22 | 632.5 | 18.4 | 1237 | US-10-473-193-4 | Sequence 4, Appl |
| 23 | 564.5 | 16.4 | 379 | US-09-813-408-11 | Sequence 11, Appl |
| 24 | 563.5 | 16.4 | 379 | US-11-156-062-23 | Sequence 23, Appl |
| 25 | 560.5 | 16.3 | 379 | US-10-202-339-2 | Sequence 2, Appl |
| 26 | 559.5 | 16.3 | 379 | US-09-920-118-14 | Sequence 14, Appl |
| 27 | 557.5 | 16.2 | 379 | US-09-813-408-10 | Sequence 10, Appl |

| | | | | | |
|----|-------|------|------|-------------------|-------------------|
| 28 | 557.5 | 16.2 | 379 | US-10-146-905A-10 | Sequence 10, Appl |
| 29 | 557.5 | 16.2 | 379 | US-10-503-706-1 | Sequence 1, Appl |
| 30 | 557.5 | 16.2 | 380 | US-10-146-905A-12 | Sequence 12, Appl |
| 31 | 556.5 | 16.2 | 379 | US-09-813-408-12 | Sequence 12, Appl |
| 32 | 554.5 | 16.1 | 372 | US-09-813-408-9 | Sequence 9, Appl |
| 33 | 551.5 | 16.0 | 378 | US-09-813-408-14 | Sequence 14, Appl |
| 34 | 550.5 | 16.0 | 379 | US-09-813-408-13 | Sequence 13, Appl |
| 35 | 547 | 15.9 | 382 | US-09-813-408-7 | Sequence 7, Appl |
| 36 | 546 | 15.9 | 380 | US-09-957-806A-10 | Sequence 10, Appl |
| 37 | 538 | 15.7 | 310 | US-09-813-408-15 | Sequence 15, Appl |
| 38 | 535.5 | 15.6 | 1398 | US-09-841-553-8 | Sequence 8, Appl |
| 39 | 535.5 | 15.6 | 1398 | US-10-090-624-6 | Sequence 6, Appl |
| 40 | 535.5 | 15.6 | 1398 | US-10-800-684-8 | Sequence 8, Appl |
| 41 | 535.5 | 15.6 | 1398 | US-10-888-588-6 | Sequence 6, Appl |
| 42 | 534.5 | 15.6 | 382 | US-10-146-905A-11 | Sequence 11, Appl |
| 43 | 533.5 | 15.5 | 382 | US-10-090-624-31 | Sequence 31, Appl |
| 44 | 533.5 | 15.5 | 382 | US-10-104-693-2 | Sequence 2, Appl |
| 45 | 533.5 | 15.5 | 382 | US-10-146-905A-8 | Sequence 8, Appl |

ALIGNMENTS

RESULT 1
US-09-841-553-1
Sequence 1, Application US/09841553
Publication No. US20020086402A1
GENERAL INFORMATION:
APPLICANT: TAKAKURA, HiKaru
MORISHITA, Mio
YANAMOTO, Katsunobu
MIRTA, Masanori
ASADA, Kiyozo
TSUNASAWA, Susumu
KATO, Ikumoshin
TITLE OF INVENTION: HYPERHYPERSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESS: Broadway and Neilmark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/841,553
FILING DATE: 24-Apr-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/894,818
FILING DATE: <Unknown>
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-Dec-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-841-553-1

Handwritten notes:
"Dis of this appl."
"Issued as 684-944H"
"Ref claim only DNA"
"No Pbl-pat."

Query Match 100.0%; Score 3437; DB 3; Length 659;
Best Local Similarity 100.0%; Pred. No. 9,3e-242;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MKRLGAVVLAIVVGLLAGTALAAPVKPVYRNAVVOQKNTGLTPGLFKYQRMNNQEV 60
DB 1 MKRLGAVVLAIVVGLLAGTALAAPVKPVYRNAVVOQKNTGLTPGLFKYQRMNNQEV 60
QY 61 DTVMFSGYDGRDRAVKVRLMGAQVRSYKIIIPAVAVKIKARDLLIAGMIDTGYGNT 120
DB 61 DTVMFSGYDGRDRAVKVRLMGAQVRSYKIIIPAVAVKIKARDLLIAGMIDTGYGNT 120
QY 121 RVSGIRFIOEDYKVOVDATSVQIGADTVNNSLGVDGQVVAIVDTGIDANHPDLKKG 180
DB 121 RVSGIRFIOEDYKVOVDATSVQIGADTVNNSLGVDGQVVAIVDTGIDANHPDLKKG 180
QY 181 VIGMYDAVNGRSTPYDDQGHGTHVAGIVAGTGSVNSQYIGVAPGAKLVGVKVLGADGSGS 240
DB 181 VIGMYDAVNGRSTPYDDQGHGTHVAGIVAGTGSVNSQYIGVAPGAKLVGVKVLGADGSGS 240
QY 241 VSTIIAGVDVYVQNKDKYKIRVINLSLSSQSSDGTDSLQAVNNAAMDAGIIVCVAAGNS 300
DB 241 VSTIIAGVDVYVQNKDKYKIRVINLSLSSQSSDGTDSLQAVNNAAMDAGIIVCVAAGNS 300
QY 301 GPNTYVGSPPAAASKVITVGAVDNDNIASFSSRGPTADGRLKEPVVAPGVDIAPRAGS 360
DB 301 GPNTYVGSPPAAASKVITVGAVDNDNIASFSSRGPTADGRLKEPVVAPGVDIAPRAGS 360
QY 361 TSMGTPINDYTTKASGTSMAITPHVSGVALILOAHSWTPDKKTKALIEIADIAPKEIA 420
DB 361 TSMGTPINDYTTKASGTSMAITPHVSGVALILOAHSWTPDKKTKALIEIADIAPKEIA 420
QY 421 DIAYGARGVNVYKAIKYDDYAKLFTGSAVDKGSATHTFDVSGATFTVATLYMDTSSDI 480
DB 421 DIAYGARGVNVYKAIKYDDYAKLFTGSAVDKGSATHTFDVSGATFTVATLYMDTSSDI 480
QY 481 DLVLYDPNGNEVDYSTAYYGFEEKVGYNPTAGTWKVVSYKGAANYQVDVSDGSLQ 540
DB 481 DLVLYDPNGNEVDYSTAYYGFEEKVGYNPTAGTWKVVSYKGAANYQVDVSDGSLQ 540
QY 541 SGGGNPNPNPNPTPTTDTQTFTGSVNDYMDTSDFTMNVNSGATKITGDLTFDTSYND 600
DB 541 SGGGNPNPNPNPTPTTDTQTFTGSVNDYMDTSDFTMNVNSGATKITGDLTFDTSYND 600
QY 601 LDLYLYDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLYASTYGMADYQLKAVVYGG 659
DB 601 LDLYLYDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLYASTYGMADYQLKAVVYGG 659
```

RESULT 2

```
US-10-090-624-12
; Sequence 12, Application US/10090624
; Publication No. US2002013235A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunobu
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
```

ORGANISM: Thermococcus celer
US-10-090-624-12

Query Match 100.0%; Score 3437; DB 4; Length 659;
Best Local Similarity 100.0%; Pred. No. 9,3e-242;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MKRLGAVVLAIVVGLLAGTALAAPVKPVYRNAVVOQKNTGLTPGLFKYQRMNNQEV 60
DB 1 MKRLGAVVLAIVVGLLAGTALAAPVKPVYRNAVVOQKNTGLTPGLFKYQRMNNQEV 60
QY 61 DTVMFSGYDGRDRAVKVRLMGAQVRSYKIIIPAVAVKIKARDLLIAGMIDTGYGNT 120
DB 61 DTVMFSGYDGRDRAVKVRLMGAQVRSYKIIIPAVAVKIKARDLLIAGMIDTGYGNT 120
QY 121 RVSGIRFIOEDYKVOVDATSVQIGADTVNNSLGVDGQVVAIVDTGIDANHPDLKKG 180
DB 121 RVSGIRFIOEDYKVOVDATSVQIGADTVNNSLGVDGQVVAIVDTGIDANHPDLKKG 180
QY 181 VIGMYDAVNGRSTPYDDQGHGTHVAGIVAGTGSVNSQYIGVAPGAKLVGVKVLGADGSGS 240
DB 181 VIGMYDAVNGRSTPYDDQGHGTHVAGIVAGTGSVNSQYIGVAPGAKLVGVKVLGADGSGS 240
QY 241 VSTIIAGVDVYVQNKDKYKIRVINLSLSSQSSDGTDSLQAVNNAAMDAGIIVCVAAGNS 300
DB 241 VSTIIAGVDVYVQNKDKYKIRVINLSLSSQSSDGTDSLQAVNNAAMDAGIIVCVAAGNS 300
QY 301 GPNTYVGSPPAAASKVITVGAVDNDNIASFSSRGPTADGRLKEPVVAPGVDIAPRAGS 360
DB 301 GPNTYVGSPPAAASKVITVGAVDNDNIASFSSRGPTADGRLKEPVVAPGVDIAPRAGS 360
QY 361 TSMGTPINDYTTKASGTSMAITPHVSGVALILOAHSWTPDKKTKALIEIADIAPKEIA 420
DB 361 TSMGTPINDYTTKASGTSMAITPHVSGVALILOAHSWTPDKKTKALIEIADIAPKEIA 420
QY 421 DIAYGARGVNVYKAIKYDDYAKLFTGSAVDKGSATHTFDVSGATFTVATLYMDTSSDI 480
DB 421 DIAYGARGVNVYKAIKYDDYAKLFTGSAVDKGSATHTFDVSGATFTVATLYMDTSSDI 480
QY 481 DLVLYDPNGNEVDYSTAYYGFEEKVGYNPTAGTWKVVSYKGAANYQVDVSDGSLQ 540
DB 481 DLVLYDPNGNEVDYSTAYYGFEEKVGYNPTAGTWKVVSYKGAANYQVDVSDGSLQ 540
QY 541 SGGGNPNPNPNPTPTTDTQTFTGSVNDYMDTSDFTMNVNSGATKITGDLTFDTSYND 600
DB 541 SGGGNPNPNPNPTPTTDTQTFTGSVNDYMDTSDFTMNVNSGATKITGDLTFDTSYND 600
QY 601 LDLYLYDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLYASTYGMADYQLKAVVYGG 659
DB 601 LDLYLYDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLYASTYGMADYQLKAVVYGG 659
```

RESULT 3

```
US-10-800-684-1
; Sequence 1, Application US/10800684
; Publication No. US20050014221A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MITTA, Masaozi
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikunobu
; TITLE OF INVENTION: ULTRATHERMOSTABLE PROTEASE GENES
; FILE REFERENCE: TAKAKURA-1A
; CURRENT APPLICATION NUMBER: US/10/800,684
; PRIOR FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: US/09/841,553
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 08/894,818
; PRIOR FILING DATE: 1997-08-29
; PRIOR APPLICATION NUMBER: JP32385/1995
```

2
This seq. no. 12, 3 & 4.
Seq. no. 12, 3 & 4.
which has not been
here.
No DH filed

PRIOR FILING DATE: 1995-12-12
 PRIOR APPLICATION NUMBER: JP96/03253
 PRIOR FILING DATE: 1996-11-07
 NUMBER OF SEQ ID NOS: 45
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 1
 LENGTH: 659
 TYPE: PR
 ORGANISM: Thermococcus celer
 US-10-800-684-1

Handwritten: This is not a claim. It is a sequence. No. 1. Not a claim.

Query Match 100.0%; Score 3437; DB 5; Length 659;
 Best Local Similarity 100.0%; Pred. No. 9.3e-242;
 Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MKRLGAVLALVYVGLAGTALAAPYKPVYRNNAVOOKYGLTTPGLFKVQRMNNQEV 60
 1 MKRLGAVLALVYVGLAGTALAAPYKPVYRNNAVOOKYGLTTPGLFKVQRMNNQEV 60
 61 DTVMGSGYGDRAVAVKRLMGAQVYKSYKII PAVAVKIKARDLLIAGMIDTGYFGNT 120
 61 DTVMGSGYGDRAVAVKRLMGAQVYKSYKII PAVAVKIKARDLLIAGMIDTGYFGNT 120
 121 RVSGIKFIEDYKQVVDATSVQIGADTYMNSLGIDSGVVAIVDTGIDANHPDLKKG 180
 121 RVSGIKFIEDYKQVVDATSVQIGADTYMNSLGIDSGVVAIVDTGIDANHPDLKKG 180
 181 VIGWYAVNAGRSPTPYDQGHGTHVAGI VAGTGSVNSQYIGVAPGAKLVGKVLGADSGS 240
 181 VIGWYAVNAGRSPTPYDQGHGTHVAGI VAGTGSVNSQYIGVAPGAKLVGKVLGADSGS 240
 241 VSTIIAGVDMVQNKDKYGRVNLISGSSQSSDGTDSLQAVNNAMDAGI VVCVAAGNS 300
 241 VSTIIAGVDMVQNKDKYGRVNLISGSSQSSDGTDSLQAVNNAMDAGI VVCVAAGNS 300
 301 GPTTYVGSPPAAASKYITVGA VDSNDNIASFSSRGPTADGRLEPVAVPQVDIIAPRAGS 360
 301 GPTTYVGSPPAAASKYITVGA VDSNDNIASFSSRGPTADGRLEPVAVPQVDIIAPRAGS 360
 361 TSMGTINDYTYKASGTSMAHPVSGVGLIIQAHPSMTDPDKYKIALIETADIAPKEIA 420
 361 TSMGTINDYTYKASGTSMAHPVSGVGLIIQAHPSMTDPDKYKIALIETADIAPKEIA 420
 421 DIAVGAARNVVAIKYDDYAKLTFGSAVDKGSATHTFVSGATFVTATLYWDTSDDI 480
 421 DIAVGAARNVVAIKYDDYAKLTFGSAVDKGSATHTFVSGATFVTATLYWDTSDDI 480
 481 DLYLDPNGNEVDYSTAYYGFEEKVGYNPTAGTWKVVSYKGAANYQVDVSDGSLSQ 540
 481 DLYLDPNGNEVDYSTAYYGFEEKVGYNPTAGTWKVVSYKGAANYQVDVSDGSLSQ 540
 541 SGGGNPNPNPNPTPTTDTOTFTGSVNDYMDTSDFTMNVNSGATKITGDLTFDTSYND 600
 541 SGGGNPNPNPNPTPTTDTOTFTGSVNDYMDTSDFTMNVNSGATKITGDLTFDTSYND 600
 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLYAVSTYGMADYQLKAVVYVG 659
 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLYAVSTYGMADYQLKAVVYVG 659

RESULT 4
 US-10-888-588-12
 Sequence 12, Application US/10888588
 Publication No. US2005084934A1
 GENERAL INFORMATION:
 APPLICANT: TAKAKURA, Hikaru
 APPLICANT: MORISHITA, Mio
 APPLICANT: SHIMOJO, Tomoko
 APPLICANT: ASADA, Kiyozo
 APPLICANT: KATO, Ikunoshin
 TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
 FILE REFERENCE: TAKAKURA-6
 CURRENT APPLICATION NUMBER: US/10/888,588

CURRENT FILING DATE: 2004-07-12
 PRIOR APPLICATION NUMBER: US/10/090,624
 PRIOR FILING DATE: 2002-03-06
 PRIOR APPLICATION NUMBER: 09/445,472
 PRIOR FILING DATE: 1999-12-06
 PRIOR APPLICATION NUMBER: 151969/1997
 PRIOR FILING DATE: 1997-06-10
 NUMBER OF SEQ ID NOS: 33
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 12
 LENGTH: 659
 TYPE: PR
 ORGANISM: Thermococcus celer
 US-10-888-588-12

Handwritten: This is not a claim. It is a sequence. No. 12. Not a claim.

Query Match 100.0%; Score 3437; DB 5; Length 659;
 Best Local Similarity 100.0%; Pred. No. 9.3e-242;
 Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MKRLGAVLALVYVGLAGTALAAPYKPVYRNNAVOOKYGLTTPGLFKVQRMNNQEV 60
 1 MKRLGAVLALVYVGLAGTALAAPYKPVYRNNAVOOKYGLTTPGLFKVQRMNNQEV 60
 61 DTVMGSGYGDRAVAVKRLMGAQVYKSYKII PAVAVKIKARDLLIAGMIDTGYFGNT 120
 61 DTVMGSGYGDRAVAVKRLMGAQVYKSYKII PAVAVKIKARDLLIAGMIDTGYFGNT 120
 121 RVSGIKFIEDYKQVVDATSVQIGADTYMNSLGIDSGVVAIVDTGIDANHPDLKKG 180
 121 RVSGIKFIEDYKQVVDATSVQIGADTYMNSLGIDSGVVAIVDTGIDANHPDLKKG 180
 181 VIGWYAVNAGRSPTPYDQGHGTHVAGI VAGTGSVNSQYIGVAPGAKLVGKVLGADSGS 240
 181 VIGWYAVNAGRSPTPYDQGHGTHVAGI VAGTGSVNSQYIGVAPGAKLVGKVLGADSGS 240
 241 VSTIIAGVDMVQNKDKYGRVNLISGSSQSSDGTDSLQAVNNAMDAGI VVCVAAGNS 300
 241 VSTIIAGVDMVQNKDKYGRVNLISGSSQSSDGTDSLQAVNNAMDAGI VVCVAAGNS 300
 301 GPTTYVGSPPAAASKYITVGA VDSNDNIASFSSRGPTADGRLEPVAVPQVDIIAPRAGS 360
 301 GPTTYVGSPPAAASKYITVGA VDSNDNIASFSSRGPTADGRLEPVAVPQVDIIAPRAGS 360
 361 TSMGTINDYTYKASGTSMAHPVSGVGLIIQAHPSMTDPDKYKIALIETADIAPKEIA 420
 361 TSMGTINDYTYKASGTSMAHPVSGVGLIIQAHPSMTDPDKYKIALIETADIAPKEIA 420
 421 DIAVGAARNVVAIKYDDYAKLTFGSAVDKGSATHTFVSGATFVTATLYWDTSDDI 480
 421 DIAVGAARNVVAIKYDDYAKLTFGSAVDKGSATHTFVSGATFVTATLYWDTSDDI 480
 481 DLYLDPNGNEVDYSTAYYGFEEKVGYNPTAGTWKVVSYKGAANYQVDVSDGSLSQ 540
 481 DLYLDPNGNEVDYSTAYYGFEEKVGYNPTAGTWKVVSYKGAANYQVDVSDGSLSQ 540
 541 SGGGNPNPNPNPTPTTDTOTFTGSVNDYMDTSDFTMNVNSGATKITGDLTFDTSYND 600
 541 SGGGNPNPNPNPTPTTDTOTFTGSVNDYMDTSDFTMNVNSGATKITGDLTFDTSYND 600
 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLYAVSTYGMADYQLKAVVYVG 659
 601 LDLYLDPNGNLVDRSTSSNSYEHVEYANPAPGTWTFVLYAVSTYGMADYQLKAVVYVG 659

RESULT 5
 US-09-841-553-5
 Sequence 5, Application US/09841553
 Publication No. US20020086402A1
 GENERAL INFORMATION:
 APPLICANT: TAKAKURA, Hikaru
 APPLICANT: MORISHITA, Mio
 APPLICANT: YAMAMOTO, Katsuniko
 APPLICANT: MITTA, Masanori


```

/ ASADA, Kiyozo
/ TSUNASAWA, Susumu
/ KATO, Ikunoshin
/ TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
/ NUMBER OF SEQUENCES: 42
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Browdy and Neimark
/ STREET: 419 Seventh Street N.W., Ste. 300
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: United States of America
/ ZIP: 20004
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/841,553
/ FILING DATE: 24-Apr-2001
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: US 08/894,818
/ FILING DATE: <Unknown>
/ APPLICATION NUMBER: JP 323285/1995
/ FILING DATE: 12-DEC-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Browdy, Roger L.
/ REGISTRATION NUMBER: 25,618
/ REFERENCE/DOCKET NUMBER: TAKAKURA=1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 628-5197
/ TELEFAX: (202) 737-3528
/
/ INFORMATION FOR SEQ ID NO: 5
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 659 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-841-553-5

Query Match      84.8%; Score 2914; DB 3; Length 659;
Best Local Similarity 83.7%; Pred. No. 1.3e-203;
Matches 558; Conservative 47; Mismatches 46; Indels 16; Gaps 7;

QY 1 MKRLGAVLALVVLGTLAALAPVKPV--VRNNAVQKRYGLITPGLFKKYQRMNMQ 58
DB 1 MKGKALILVLVLGLVGVSAAPBEKKVQVRN---VEKNYGLITPGLFKKIQKLNPN 57
QY 59 EVDVIMFGSGDRRAVKVRLMGAVKYSYKIIPAVAVKIKARDLLIAGMIDTGFG 118
DB 58 EISTIVIEENHREKEIAVRVLELMGAKRVYVHIIPALADLKVRDLVLSGL--TG--G 113
QY 119 NTRVSGIKFTQEDYKVOYDVA-----TSVSGIGADTYWNSLGYDGSVVVAIVDTGIDAN 173
DB 114 KAKLSGVRFIQEDYKVTVSABEGLSAQAQWATYWN--LGIDSGSITIGITDGLDAS 172
QY 174 HPDLKGVITGVYDAVNGRSTPYDDQGHGTHVAGIVAGTGSV-NSQYTGAVPAGAKLVGVK 232
DB 173 HPDLQGVITGVYDAVNGRSTPYDDHGHGTHVASIAAGTGAASNGKYGAPAKLAGITKV 232
QY 233 LGADSGSVSTIIAGVDVWVQNDKYGIRYINSLSSSOSDGTDSLQAVNNAMAGIV 292
DB 233 LGADSGSISTIIKGVMAVDNDKGIKVINSLSSSOSDGTDSLQAVNNAMAGIV 292
QY 293 VCVAAAGNSGNTYVGSPPAASKVITVGAVDSDNDIASFSSRGTADGRLKEPVAPGV 352
DB 293 VCVAAAGNSGNTYVGSPPAASKVITVGAVDSDNDIASFSSRGTADGRLKEPVAPGV 352
QY 353 IIAPRASGTSMPINDYTTKASGTSMAIPTHVSGVALLIOAHPSWTPDKVKTALLETAD 412
DB 353 IIAPRASGTSMPINDYTTKASGTSMAIPTHVSGVALLIOAHPSWTPDKVKTALLETAD 412

```

```

QY 413 IVAPKEIADTAYGAGRVNVTIKIKYDDYAKLTFTGGSVADKGSATHTFDVSGATFVATLLY 472
DB 413 IVAPKEIADTAYGAGRVNVTIKIKYDDYAKLTFTGGSVADKGSATHTFDVSGATFVATLLY 472
QY 473 WDTSDDIDILYLDVDPNGNEVDYSYTAAYGFEKGYVNPAGTMTVKKVSKGAANYQVDV 532
DB 473 WDTSDDIDILYLDVDPNGNEVDYSYTAAYGFEKGYVNPAGTMTVKKVSKGAANYQVDV 532
QY 533 VSDGSLSGSGGQNPENPNPTPTDQTQFTGSVNDYWDTSDFTMVNVSGATKITGDL 592
DB 533 VSDGSLSGSGGQNPENPNPTPTDQTQFTGSVNDYWDTSDFTMVNVSGATKITGDL 592
QY 593 TFPDSVNDLILYLDVDPNGNLVDRSTSSNSYEHTYANPAPGTTWPLVYASTGADYQL 652
DB 593 TFPDSVNDLILYLDVDPNGNLVDRSTSSNSYEHTYANPAPGTTWPLVYASTGADYQL 652
QY 653 KAVVYVG 659
DB 653 KAVVYVG 659

RESULT 6
US-10-800-684-5
/ Sequence 5, Application US/10800684
/ Publication No. US20050014221A1
/ GENERAL INFORMATION:
/ APPLICANT: TAKAKURA, Hikaru
/ APPLICANT: MORISHITA, Mio
/ APPLICANT: YAMAMOTO, Katsuhiko
/ APPLICANT: MITTA, Masanori
/ APPLICANT: ASADA, Kiyozo
/ APPLICANT: TSUNASAWA, Susumu
/ APPLICANT: KATO, Ikunoshin
/
/ TITLE OF INVENTION: ULTRATHERMOSTABLE PROTEASE GENES
/ FILE REFERENCE: TAKAKURA=1A
/ CURRENT APPLICATION NUMBER: US/10/800,684
/ PRIOR FILING DATE: 2004-03-16
/ PRIOR APPLICATION NUMBER: US/09/841,553
/ PRIOR FILING DATE: 2001-04-24
/ PRIOR APPLICATION NUMBER: 08/894,818
/ PRIOR FILING DATE: 1997-08-29
/ PRIOR APPLICATION NUMBER: JP32385/1995
/ PRIOR FILING DATE: 1995-12-12
/ PRIOR APPLICATION NUMBER: JP96/03253
/ PRIOR FILING DATE: 1996-11-07
/ NUMBER OF SEQ ID NOS: 45
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 5
/ LENGTH: 659
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic
US-10-800-684-5

Query Match      84.8%; Score 2914; DB 5; Length 659;
Best Local Similarity 83.7%; Pred. No. 1.3e-203;
Matches 558; Conservative 47; Mismatches 46; Indels 16; Gaps 7;

QY 1 MKRLGAVLALVVLGTLAALAPVKPV--VRNNAVQKRYGLITPGLFKKYQRMNMQ 58
DB 1 MKGKALILVLVLGLVGVSAAPBEKKVQVRN---VEKNYGLITPGLFKKIQKLNPN 57
QY 59 EVDVIMFGSGDRRAVKVRLMGAVKYSYKIIPAVAVKIKARDLLIAGMIDTGFG 118
DB 58 EISTIVIEENHREKEIAVRVLELMGAKRVYVHIIPALADLKVRDLVLSGL--TG--G 113
QY 119 NTRVSGIKFTQEDYKVOYDVA-----TSVSGIGADTYWNSLGYDGSVVVAIVDTGIDAN 173
DB 114 KAKLSGVRFIQEDYKVTVSABEGLSAQAQWATYWN--LGIDSGSITIGITDGLDAS 172
QY 174 HPDLKGVITGVYDAVNGRSTPYDDQGHGTHVAGIVAGTGSV-NSQYTGAVPAGAKLVGVK 232

```



```

Db 173 HPDLQKVLGWDFVNGRSYFYDDHGHGTHVASIAAGTGAASNGKKGMAFGAKLAGIKY 232
Qy 233 LGADSGSVSTIIAGVDVWQNDKXGIRYINLSLSSOSSDCTDLSQAVNNAMAGIV 292
Db 233 LGADSGSVSTIIAGVDVWQNDKXGIRYINLSLSSOSSDCTDLSQAVNNAMAGIV 292
Qy 293 VCVAAGNSGPNITYTGSPPAASKVIITVGAVDSNDNIASFSSRGPTADGRLKEPVAPGV 352
Db 293 VCVAAGNSGPNITYTGSPPAASKVIITVGAVDSNDNIASFSSRGPTADGRLKEPVAPGV 352
Qy 353 IIAPRASGSTMGPINDYTTKASGTMATPHVSGVALLIOAHPSWTPDKVTALLETAD 412
Db 353 IIAPRASGSTMGPINDYTTKASGTMATPHVSGVALLIOAHPSWTPDKVTALLETAD 412
Qy 413 IYAPKEIADIAVAGAVNNYKAIKYDDYAKLTFTGSVADKGSATHTFVSGATFVATLY 472
Db 413 IYAPKEIADIAVAGAVNNYKAIKYDDYAKLTFTGSVADKGSATHTFVSGATFVATLY 472
Qy 473 WDTGSSDIDLXYDPNGNEVDYSYTAAYGFEKGYNNPTAGTWTVKVSYKGAANYQVDY 532
Db 473 WDTGSSDIDLXYDPNGNEVDYSYTAAYGFEKGYNNPTAGTWTVKVSYKGAANYQVDY 532
Qy 533 VSDGSLSSGCGGNPNPNPPTTTDTOTFTGSVANDYDTSPTTMANNSGATKITGDL 592
Db 533 VSDGSLSSGCGGNPNPNPPTTTDTOTFTGSVANDYDTSPTTMANNSGATKITGDL 592
Qy 593 TPDTSNDIDLXYDPNGNLVDRSTSSNSYEHVEYANPAPGTWFLVAYSTYGADYOL 652
Db 593 TPDTSNDIDLXYDPNGNLVDRSTSSNSYEHVEYANPAPGTWFLVAYSTYGADYOL 652
Qy 653 KAVVYTG 659
Db 653 KAVVYTG 659

```

RESULT 7

US-09-841-553-35

Sequence 35, Application US/09841553
 Publication No. US20020086402A1

GENERAL INFORMATION:

APPLICANT: TAKAKURA, Hikaru

YAMAMOTO, Katsuhiko

MORISHITA, Mio

MITTA, Masanori

ASADA, Kiyozo

TSUNASAWA, Susumu

KATO, Ikunoshin

TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESS: Browdy and Neimark

STREET: 419 Seventh Street N.W., Ste. 300

CITY: Washington

STATE: D.C.

COUNTRY: United States of America

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/841,553

FILING DATE: 24-Apr-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/894,818

FILING DATE: <Unknown>

APPLICATION NUMBER: JP 323285/1995

FILING DATE: 12-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Browdy, Roger L.

REGISTRATION NUMBER: 25,618

```

REFERENCE/DOCKET NUMBER: TAKAKURA=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-841-553-35

```

Query Match 72.3%; Score 2483.5; DB 3; Length 654;
 Best Local Similarity 72.1%; Pred. No. 3 2e-172;
 Matches 461; Conservative 69; Mismatches 96; Indels 21; Gaps 8;

```

Qy 1 MRLGAVNLALVGLAGTALAAPYKPV--VRNNAVQOKNTGLTPGLFKVQRMNMQ 58
Db 1 MGLKALILVILVGLVGSVAAPKVEQVRN--VEKNYGLTPGLFRKIQKLNPN 57
Qy 59 EVDYVMEGSGYDRDRAVYFLMGAQVYSKIIIPAVAVKIKARULLIAGHIDGYRG 118
Db 58 EISTVIVFENHREKEIAVFLIMGAKRYVYHIIIPAIADIKVRLDVISGL--TG--G 113
Qy 119 NTRVSGIKFIOEDYKQVQVDA-----TSVSOIGADTVNNSLGVDGCVVAIVDTGIDAN 173
Db 114 KAKLSGVRFIOEDYKQVYSAELBGLDESAQWATYVNN--LGDSGIGITGIIDTIDAS 172
Qy 174 HPDLQKVLGWDAVNGRSTPYDDOGHGTHVAGIVAGTGSV--NSQYIGVAPGAKLVGVY 232
Db 173 HPDLQKVLGWDFVNGRSYFYDDHGHGTHVASIAAGTGAASNGKKGMAFGAKLAGIKY 232
Qy 233 LGADSGSVSTIIAGVDVWQNDKXGIRYINLSLSSOSSDCTDLSQAVNNAMAGIV 292
Db 233 LGADSGSVSTIIAGVDVWQNDKXGIRYINLSLSSOSSDCTDLSQAVNNAMAGIV 292
Qy 293 VCVAAGNSGPNITYTGSPPAASKVIITVGAVDSNDNIASFSSRGPTADGRLKEPVAPGV 352
Db 293 VCVAAGNSGPNITYTGSPPAASKVIITVGAVDSNDNIASFSSRGPTADGRLKEPVAPGV 352
Qy 353 IIAPRASGSTMGPINDYTTKASGTMATPHVSGVALLIOAHPSWTPDKVTALLETAD 412
Db 353 IIAPRASGSTMGPINDYTTKASGTMATPHVSGVALLIOAHPSWTPDKVTALLETAD 412
Qy 413 IYAPKEIADIAVAGAVNNYKAIKYDDYAKLTFTGSVADKGSATHTFVSGATFVATLY 472
Db 413 IYAPKEIADIAVAGAVNNYKAIKYDDYAKLTFTGSVADKGSATHTFVSGATFVATLY 472
Qy 473 WDTGSSDIDLXYDPNGNEVDYSYTAAYGFEKGYNNPTAGTWTVKVSYKGAANYQVDY 532
Db 473 WDTGSSDIDLXYDPNGNEVDYSYTAAYGFEKGYNNPTAGTWTVKVSYKGAANYQVDY 532
Qy 533 VSDGSLSSGCGGNPNPNPPTTTDTOTFTGSVANDYDTSPTTMANNSGATKITGDL 592
Db 533 VSDGSLSSGCGGNPNPNPPTTTDTOTFTGSVANDYDTSPTTMANNSGATKITGDL 592
Qy 593 TPDTSNDIDLXYDPNGNLVDRSTSSNSYEHVEYANPAPGTWFLVAYSTYGADYOL 652
Db 588 VPDTSYHDDIDLXYDPNGNLVDRSESPNSYEHVEYLTAPAGTWFLVAYSTYGADYOL 647
Qy 653 KAVVYTG 659
Db 648 TAKVYTG 654

```

RESULT 8

US-10-090-624-16

Sequence 16, Application US/10090624

Publication No. US20020132335A1

GENERAL INFORMATION:

APPLICANT: TAKAKURA, Hikaru

APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOJO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/10/090,624
CURRENT FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: 09/445,472
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 16
LENGTH: 654
TYPE: PRF
ORGANISM: Pyrococcus furiosus
US-10-090-624-16

Query Match 72.3%; Score 2483.5; DB 4; Length 654;
Best Local Similarity 72.1%; Pred. No. 3.2e-172;
Matches 481; Conservative 69; Mismatches 96; Indels 21; Gaps 8;

1 MKRLGAVLALVVLVGLAGTALAAPVKPV--VRNNAVQCKNYGLTPGLFKYQRMNMNQ 58
1 MKGKALLLVILVIGLVGSVAAPBEKVEQVRN---VEKNYGLTGLFKIKLPNE 57
59 EVDVTVMPSGYDRDAVAVLRLMGAQVYKSIIPAVAVKIKARLLIAGMTDGYFG 118
58 EISTVIVFNHREKELAVLVELMGAKRYVYHIIIPAIADLKVDLVIISGL--TG--G 113
119 NTRVSGIKFIEDYKQVODA-----TSVQIGADTVNNSLGYSQGVVAIVDTGIDAN 173
114 KAKLSGVRFIEDYKVTVAELGLDESAQVMATYVM--LGYDGSGITIGIIDTGIDAS 172
174 HPDLKGVIGWYDAVNGRSTPYDDQGHGTHVAGIVAGTGSV--NSQYIGVAPAKLVGVKV 232
173 HPDLQGVIGWYDAVNGRSTPYDDHGHGTHVASIAAGTGAANGKYGMAPGAALAGIKV 232
233 LGADGSGSVSTIIAGVDVAVVONKDKYGRVNLISGSSQSDGTDSLQAVNNAMDAGIV 292
233 LGADGSGSVSTIIAGVEMAVDNKDKYGRVNLISGSSQSDGTDSLQAVNNAMDAGIV 292
293 VCVAAAGNSGPNNTYVGSPPAAASKVITVGAVDSDNINASFSSRGPTADGRLKEPVVAPGV 352
293 VVVAAGNSGPNNTYVGSPPAAASKVITVGAVDKYDVTISFSSRGPTADGRLKEPVVAPGV 352
353 IIAPRASGSMGPTIINDYTTKASGSMATPHVSGVGLIIQAHPSWTPDKVKTALILETAD 412
353 IIAPRASGSMGPTIINDYTTKASGSMATPHVAGIALLIQAHPSPWTPDKVKTALILETAD 412
413 IYAPKEIADIAYGAGRVNYYKAIKYDDYAKLFTGVSADKGSATHTPDVSGATFVTATLY 472
413 IYKPEIADIAYGAGRVNYYKAIKYDDYAKLFTGVSADKGSATHTPDVSGATFVTATLY 472
473 WDTGSSSIDLYLIDPNNGEVDSYTAAYGFEKGYGYNPTAGTWTVKVSYKGAANYQVDV 532
473 WDNANSDDLILYLDPNNGQVDYSYTAAYGFEKGYGYNPTAGTWTIKVSYSGSANYQVDV 532
533 VSDGSLSGSGGNPNPNPNPTTTDTOTFTGTSVNDYMDTSDPTMNVNNSGATKITGDL 592
533 VSDGSLSGSGGNPNPNPNPTTTDTOTFTGTSVNDYMDTSDPTMNVNNSGATKITGDL 592
533 VSDGSLSGSGGNPNPNPNPTTTDTOTFTGTSVNDYMDTSDPTMNVNNSGATKITGDL 592
593 TPTSTYNDLILYLDPNNGEVDSYTAAYGFEKGYGYNPTAGTWTVKVSYKGAANYQVDV 532
593 TPTSTYNDLILYLDPNNGEVDSYTAAYGFEKGYGYNPTAGTWTVKVSYKGAANYQVDV 532
588 VFDTSYNDLILYLDPNNGEVDSYTAAYGFEKGYGYNPTAGTWTIKVSYSGSANYQVDV 532
653 KAVVYVG 659
648 YAKVYVG 654

RESULT 9
US-10-800-684-35
Sequence 35, Application US/10800684
Publication No. US20050014221A1
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MITTA, Masanori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
TITLE OF INVENTION: ULTRATHERMOSTABLE PROTEASE GENES
FILE REFERENCE: TAKAKURA-1A
CURRENT APPLICATION NUMBER: US/10/800,684
CURRENT FILING DATE: 2004-03-16
PRIOR APPLICATION NUMBER: US/09/841,553
PRIOR FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: 08/894,818
PRIOR FILING DATE: 1997-08-29
PRIOR APPLICATION NUMBER: JP32385/1995
PRIOR FILING DATE: 1995-12-12
PRIOR APPLICATION NUMBER: JP96/03253
PRIOR FILING DATE: 1996-11-07
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn version 3.2
SEQ ID NO 35
LENGTH: 654
TYPE: PRF
ORGANISM: Pyrococcus furiosus
US-10-800-684-35

Query Match 72.3%; Score 2483.5; DB 5; Length 654;
Best Local Similarity 72.1%; Pred. No. 3.2e-172;
Matches 481; Conservative 69; Mismatches 96; Indels 21; Gaps 8;

1 MKRLGAVLALVVLVGLAGTALAAPVKPV--VRNNAVQCKNYGLTPGLFKYQRMNMNQ 58
1 MKGKALLLVILVIGLVGSVAAPBEKVEQVRN---VEKNYGLTGLFKIKLPNE 57
59 EVDVTVMPSGYDRDAVAVLRLMGAQVYKSIIPAVAVKIKARLLIAGMTDGYFG 118
58 EISTVIVFNHREKELAVLVELMGAKRYVYHIIIPAIADLKVDLVIISGL--TG--G 113
119 NTRVSGIKFIEDYKQVODA-----TSVQIGADTVNNSLGYSQGVVAIVDTGIDAN 173
114 KAKLSGVRFIEDYKVTVAELGLDESAQVMATYVM--LGYDGSGITIGIIDTGIDAS 172
174 HPDLKGVIGWYDAVNGRSTPYDDQGHGTHVAGIVAGTGSV--NSQYIGVAPAKLVGVKV 232
173 HPDLQGVIGWYDAVNGRSTPYDDHGHGTHVASIAAGTGAANGKYGMAPGAALAGIKV 232
233 LGADGSGSVSTIIAGVDVAVVONKDKYGRVNLISGSSQSDGTDSLQAVNNAMDAGIV 292
233 LGADGSGSVSTIIAGVEMAVDNKDKYGRVNLISGSSQSDGTDSLQAVNNAMDAGIV 292
293 VCVAAAGNSGPNNTYVGSPPAAASKVITVGAVDSDNINASFSSRGPTADGRLKEPVVAPGV 352
293 VVVAAGNSGPNNTYVGSPPAAASKVITVGAVDKYDVTISFSSRGPTADGRLKEPVVAPGV 352
353 IIAPRASGSMGPTIINDYTTKASGSMATPHVSGVGLIIQAHPSWTPDKVKTALILETAD 412
353 IIAPRASGSMGPTIINDYTTKASGSMATPHVAGIALLIQAHPSPWTPDKVKTALILETAD 412
413 IYAPKEIADIAYGAGRVNYYKAIKYDDYAKLFTGVSADKGSATHTPDVSGATFVTATLY 472
413 IYKPEIADIAYGAGRVNYYKAIKYDDYAKLFTGVSADKGSATHTPDVSGATFVTATLY 472
473 WDTGSSSIDLYLIDPNNGEVDSYTAAYGFEKGYGYNPTAGTWTVKVSYKGAANYQVDV 532
473 WDNANSDDLILYLDPNNGQVDYSYTAAYGFEKGYGYNPTAGTWTIKVSYSGSANYQVDV 532
533 VSDGSLSGSGGNPNPNPNPTTTDTOTFTGTSVNDYMDTSDPTMNVNNSGATKITGDL 592

| | | | | |
|---------------------------|-------|---------------------|-----------|-------------|
| Query Match | 62.2% | Score 2138.5; | DB 3; | Length 522; |
| Best Local Similarity | 78.3% | Pred. No. 3.4e-147; | | |
| Matches 407; Conservative | 36; | Mismatches 70; | Indels 7; | Gaps 3; |

```

QY 141 SVSOGIDADTWNNSLGYGSGSVVAIVADTGDIDANHPDLKSKYIGWDAVNGSTPYDQGH 200
Db 9 SAAQVMAITYWN-LGYDSGSGITIGIIDTGDASHPDLQSKYIGWDPFVNGSGSPYDHDGH 67
QY 201 GTHVAGIYAIGTSV-NSQYIGVAPGAKLGVCKYLGDSSGSVSTIIAGVDWVYONKRYG 255
Db 68 GTHVAAIAAGGAASNKKYMGAPAKLAGIKVLGADGSGSISTIIKGVENAVONKRYG 127
QY 260 IRVINTLSGSSOSBDGTDLSQAVNNAMAGIYVCAAGNSGPNITYGSPRAASKYITV 319
Db 128 IKVINTLSGSSGSDGTDLSQAVNNAMAGIYVVAAGNSGPNKTYIGSPRAASKYITV 187
QY 320 GAVDSNDNINASFSSRGFTADGRLKEPEVAPGVDIAPRASGTSWGPINDYYTYRASGTSM 379
Db 188 GAVDKXDVINTSFSSRGFTADGRLKEPEVAPGAMWIIAARASGTSWQOPINDYYTYAAPGISM 247
QY 380 ATPVSGVGLIIIOAHPSMTPOKVKKALLETADIYVAPKEIADINGAARVNVYAKIKYD 439
Db 248 ATPHVAGIALLIOAHPSMTPOKVKKALLETADIYKPEIADIAGAARVAAYAIAYND 307
QY 440 YAKLTFTGSVADKXSAHTTFDVSAGATFVATLLYMDTGSSDIDLVLDPNGENIEVDYSYAY 499
Db 308 YAKLVFTGVYANKSGSQHQGFVJISGASFVATLLYMDNANSDDLVLDPNGQVQVSYTAY 367
QY 500 YGPEKXUYVNPFLAGTITVNVKYSKGAANQVUDVVSDDLSGSGGGENPNEPNPPTTDD 559
Db 368 YGPEKXGYVNPFDGTTITKIVKYSISGSAANQVUDVVSDDGISQ-----PSSSPSPQEPETVD 422
QY 560 TQTFGSVNDYMDTSDTFTMANVSGATIKTGDLTFTDSYNDLIDLVLDPNKNLVDRTSS 619
Db 423 AKTFQXSDHYHYDRSDTFTMTVNSGATIKTGDLVFTDSYHDDLVLDPNKNLYDRBSRP 482
QY 620 NSYEHVEYANPAGTITFLVYAYSTYGMADVQAKAVVYTG 659
Db 483 NSYEHVEYLPAGTITFLVYAYYTGAAVYELTAKVYTG 522

```

RESULT 12

```

US-10-090-624-4
/ Sequence 4, Application US/10090624
/ Publication No. US2002013335A1
/ GENERAL INFORMATION:
/ APPLICANT: TAKAKURA, Hikaru
/ APPLICANT: MORISHITA, Mio
/ APPLICANT: SHIMOMO, Tomoko
/ APPLICANT: ASADA, Kiyozo
/ APPLICANT: KATO, Ikunoshih
/ TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
/ FILE REFERENCE: TAKAKURA-6
/ CURRENT APPLICATION NUMBER: US/10/090,624
/ CURRENT FILING DATE: 2002-03-06
/ PRIOR APPLICATION NUMBER: 09/445,472
/ PRIOR FILING DATE: 1999-12-06
/ PRIOR APPLICATION NUMBER: 151969/1997
/ PRIOR FILING DATE: 1997-06-10
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 4
/ LENGTH: 522
/ TYPE: PR1
/ ORGANISM: Pyrococcus furiosus
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (428)..(428)
/ OTHER INFORMATION: Xaa at position 428 is Gly or Val.
US-10-090-624-4

```

Best Local Similarity 78.3%; Pred. NO. 3.4e-147;
Matches 407; Conservative 36; Mismatches 70; Indels 7; Gaps 3;

[illegible]

RESULT 13

```

US-10-800-684-3
Sequence 3, Application US/10800684
Publication No. US20050014221A1
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hi-karu
APPLICANT: MORISHITA, Mitsuhiro
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MITTA, Masanori
APPLICANT: ASADA, Kiyo-ozo
APPLICANT: TSUNASAWA, Susumu
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: ULTRAHERMOSTABLE PROTEASE GENES
FILE REFERENCE: TAKAKURA-1A
CURRENT APPLICATION NUMBER: US/10/800,684
CURRENT FILING DATE: 2004-03-16
PRIOR APPLICATION NUMBER: US/09/841,553
PRIOR FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: 08/894,818
PRIOR FILING DATE: 1997-08-29
PRIOR APPLICATION NUMBER: JP32365/1995
PRIOR FILING DATE: 1995-12-13
PRIOR APPLICATION NUMBER: JP96/03253
PRIOR FILING DATE: 1996-11-07
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3
LENGTH: 522
TYPE: PRT
ORGANISM: Pyrococcus furiosus
FEATURE:
NAME/KEY: misc_feature
LOCATION: (428)..(428)

```

OTHER INFORMATION: Xaa is Gly or Val
US-10-800-684-3

Query Match 62.2%; Score 2138.5; DB 5; Length 522;

Best Local Similarity 78.3%; Pred. No. 3.4e-147;
Matches 407; Conservative 36; Mismatches 70; Indels 7; Gaps 3;

```
QY 141 SVSQTGADTVNMSLGDGSGVVAIVDTGIDANHPDLKGVIGWYDAVNGRSTPYDDQH 200
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 SAAQWATYVNM-LGVDGSGITIGITIDGIDASHPLDQKVGWDFVNGRSTPYDDQH 67
QY 201 GTHVAGIYAGTGSV-NSQYIGVAPGAKLVGVKYLGDGSGSVSTIIAGVDMVQNKDKTG 259
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 GTHVASIAAGTGAASNGKYGWAPGAKLAGIKVLGDGSGSISTIIKGVEMAVDNKDKYG 127
QY 260 IRVINLSLSSSOSSGTSDLSQAVNNAADAGIVCAAGNSGPNMTYVCSPPAAASKVITV 319
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 IKVINLSLSSSOSSGTDLSQLVNAAMDAGLVVVAAGNSGPNKTYTIGSPAAASKVITV 187
QY 320 GAVDSNDNIASFSSRGPTADGRLKPEVVA PGVDIIAPRASGTSMTGPIINDYTTKASGTSM 379
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 GAVDKVDVITSSSRGPTADGRLKPEVVA PGWIIIAARASGTSMTGPIINDYTTAARGTSM 247
QY 380 ATPHVSQVGAALILQAHPSWTPDKVTALLETADIAPKEIADIAGAGRVNYYKAIKYDD 439
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 ATPHVAIGIALLLQAHPSWTPDKVTALLETADIAPKEIADIAGAGRVNYYKAIKYDD 307
QY 440 YAKLFTGVSADKSGATHTFDVSGATFPTATLYMTGSSDIDLXYDPNGNEVDVSYTAY 499
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 308 YAKLVFTGVANKGQTHQFVLSGAS FVTATLYMNANSDDLXYDPNGNOVDVSYTAY 367
QY 500 YGEFEKGYNPPAGTWTWYKVSYSKGAANYQVDVSDGSLSSGSGGPNPNPNPTPTD 559
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 368 YGEFEKGYNPPDGWTITIKVSYSGSANYQVDVSDGSLSSGSGSPPEPTD 422
QY 560 TOTFTGSVNDWDTSDFTTMVNSGATKITGDLTFEDTSYNDLXYDPNGNLVDRSTSS 619
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 423 AKTFQXSDHYYYDRSDFTMTVNSGATKITGDLVFDTSYHDDLXYDPNGNLVDRSESP 482
QY 620 NSYEHEVYANPAPGWTFLVYAYSTYGWADYOLKAVYVYG 659
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 483 NSYEHEVYLTTPAGTMYFLVYAYTYGWAYYELTAKVYVG 522
```

RESULT 14
US-10-888-588-4

Sequence 4, Application US/10888588
Publication No. US20050084934A1

GENERAL INFORMATION:

APPLICANT: MORISHITA, Hikaru
APPLICANT: SHIMODO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/10/888,588
CURRENT FILING DATE: 2004-07-12
PRIOR APPLICATION NUMBER: US/10/090,624
PRIOR FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: 09/445,472
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0

SEQ ID NO 4

LENGTH: 522

TYPE: PRT

ORGANISM: Pyrococcus furiosus

FEATURE:

NAME/KEY: misc_feature
LOCATION: (428)..(428)

OTHER INFORMATION: Xaa at position 428 is Gly or Val.
US-10-888-588-4

Query Match 62.2%; Score 2138.5; DB 5; Length 522;

Best Local Similarity 78.3%; Pred. No. 3.4e-147;
Matches 407; Conservative 36; Mismatches 70; Indels 7; Gaps 3;

```
QY 141 SVSQTGADTVNMSLGDGSGVVAIVDTGIDANHPDLKGVIGWYDAVNGRSTPYDDQH 200
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 SAAQWATYVNM-LGVDGSGITIGITIDGIDASHPLDQKVGWDFVNGRSTPYDDQH 67
QY 201 GTHVAGIYAGTGSV-NSQYIGVAPGAKLVGVKYLGDGSGSVSTIIAGVDMVQNKDKTG 259
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 GTHVASIAAGTGAASNGKYGWAPGAKLAGIKVLGDGSGSISTIIKGVEMAVDNKDKYG 127
QY 260 IRVINLSLSSSOSSGTSDLSQAVNNAADAGIVCAAGNSGPNMTYVCSPPAAASKVITV 319
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 IKVINLSLSSSOSSGTDLSQLVNAAMDAGLVVVAAGNSGPNKTYTIGSPAAASKVITV 187
QY 320 GAVDSNDNIASFSSRGPTADGRLKPEVVA PGVDIIAPRASGTSMTGPIINDYTTKASGTSM 379
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 GAVDKVDVITSSSRGPTADGRLKPEVVA PGWIIIAARASGTSMTGPIINDYTTAARGTSM 247
QY 380 ATPHVSQVGAALILQAHPSWTPDKVTALLETADIAPKEIADIAGAGRVNYYKAIKYDD 439
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 ATPHVAIGIALLLQAHPSWTPDKVTALLETADIAPKEIADIAGAGRVNYYKAIKYDD 307
QY 440 YAKLFTGVSADKSGATHTFDVSGATFPTATLYMTGSSDIDLXYDPNGNEVDVSYTAY 499
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 308 YAKLVFTGVANKGQTHQFVLSGAS FVTATLYMNANSDDLXYDPNGNOVDVSYTAY 367
QY 500 YGEFEKGYNPPAGTWTWYKVSYSKGAANYQVDVSDGSLSSGSGGPNPNPNPTPTD 559
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 368 YGEFEKGYNPPDGWTITIKVSYSGSANYQVDVSDGSLSSGSGSPPEPTD 422
QY 560 TOTFTGSVNDWDTSDFTTMVNSGATKITGDLTFEDTSYNDLXYDPNGNLVDRSTSS 619
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 423 AKTFQXSDHYYYDRSDFTMTVNSGATKITGDLVFDTSYHDDLXYDPNGNLVDRSESP 482
QY 620 NSYEHEVYANPAPGWTFLVYAYSTYGWADYOLKAVYVYG 659
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 483 NSYEHEVYLTTPAGTMYFLVYAYTYGWAYYELTAKVYVG 522
```

RESULT 15
US-10-090-624-1

Sequence 1, Application US/10090624
Publication No. US2002013235A1

GENERAL INFORMATION:

APPLICANT: MORISHITA, Hikaru
APPLICANT: SHIMODO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/10/090,624
CURRENT FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: 09/445,472
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0

SEQ ID NO 1

LENGTH: 412

TYPE: PRT

ORGANISM: Pyrococcus furiosus

FEATURE:

NAME/KEY: misc_feature
US-10-090-624-1

Query Match 49.7%; Score 1707; DB 4; Length 412;
Best Local Similarity 80.6%; Pred. No. 7.3e-116;
Matches 325; Conservative 29; Mismatches 47; Indels 2; Gaps 2;

QY 141 SVEQIGADTVWNSIGYDGSVVVAIVDTGIDANHPDLKGVIGWYDAVNGRSTPYDDQGH 200
 Db 9 SAAQVMAITYVMN-LGYDGSGITIGIDTGDASHPDLQGVIGWYDAVNGRSTPYDDHGH 67
 QY 201 GTHVAGIVAGTGV-NSQYIGVAPGAKLVGVKYLADGSGSVSTIIAGVDMVQNKDKYG 259
 Db 68 GTHVASIAGTGAASNGKYGMAFGAKLAGIKVLAGDGSISITIIKGVEMAVDNKDKYG 127
 QY 260 IRVINISLGSOSDGTDSLQAVNNAMDAIVVCVAGNSGPNYTVGS PAAASKVITY 319
 Db 128 IKVINISLGSOSDGTDSLQAVNNAMDAIVVCVAGNSGPNYTVGS PAAASKVITY 187
 QY 320 GAVDSNDNIASFSSRGPTADGRLKPEVAPGVDIIPRASGSMGTPINDYTKASGTSM 379
 Db 188 GAVDKYDVITSFSSRGPTADGRLKPEVAPGVDIIPRASGSMGTPINDYTKASGTSM 247
 QY 380 ATPHVSQVGLIIOAPSWTPDKYKTLIETADIAPKEIADIAYGAGRNVYKAIKYDD 439
 Db 248 ATPHVAGIALLLQAPSWTPDKYKTLIETADIVKPEIADIAYGAGRNVYKAIKYDD 307
 QY 440 YAKLFTGSAVDKGSATHTFDVSGATPVTATLYMDTGSSDIDLXYDPNGNEVDYSYAY 499
 Db 308 YAKLVFTGYVANKSSQTHQFVIGASFTVATLYMDNANSDDLXYDPNGNEVDYSYAY 367
 QY 500 YGEKVGYNPTAGTWTVKVVSYKGAANYQVDVSDGSLQSG 542
 Db 368 YGEKVGYNPTAGTWTIKVVSYSGSANYQVDVSDGSLQSG 410

Search completed: January 6, 2007, 22:37:52
 Job time : 124.5 secs